

# Introduction to Transformational Grammar

Kyle Johnson  
University of Massachusetts at Amherst

Fall 2004



# Contents

<b>Preface</b>	<b>iii</b>
<b>1 The Subject Matter</b>	<b>1</b>
1.1 Linguistics as learning theory . . . . .	1
1.2 The evidential basis of syntactic theory . . . . .	7
<b>2 Phrase Structure</b>	<b>15</b>
2.1 Substitution Classes . . . . .	16
2.2 Phrases . . . . .	20
2.3 $\bar{X}$ phrases . . . . .	29
2.4 Arguments and Modifiers . . . . .	41
<b>3 Positioning Arguments</b>	<b>57</b>
3.1 Expletives and the Extended Projection Principle . . . . .	58
3.2 Case Theory and ordering complements . . . . .	61
3.3 Small Clauses and the Derived Subjects Hypothesis . . . . .	68
3.4 PRO and Control Infinitives . . . . .	79
3.5 Evidence for Argument Movement from Quantifier Float . . . . .	83
3.6 Towards a typology of infinitive types . . . . .	92
3.7 Constraints on Argument Movement and the typology of verbs . . . . .	97
<b>4 Verb Movement</b>	<b>105</b>
4.1 The “Classic” Verb Movement account . . . . .	106
4.2 Head Movement’s role in “Verb Second” word order . . . . .	115
4.3 The Pollockian revolution: exploded IPs . . . . .	123
4.4 Features and covert movement . . . . .	136
<b>5 Determiner Phrases and Noun Movement</b>	<b>149</b>
5.1 The DP Hypothesis . . . . .	151
5.2 Noun Movement . . . . .	155

## Contents

---

<b>6</b>	<b>Complement Structure</b>	<b>179</b>
6.1	Nouns and the $\theta$ -roles they assign . . . . .	180
6.2	Double Object constructions and Larsonian shells . . . . .	195
6.3	Complement structure and Object Shift . . . . .	207
<b>7</b>	<b>Subjects and Complex Predicates</b>	<b>229</b>
7.1	Getting into the right position . . . . .	229
7.2	Subject Arguments . . . . .	233
7.2.1	Argument Structure . . . . .	235
7.2.2	The syntactic benefits of $v$ . . . . .	245
7.3	The relative positions of $\mu P$ and $vP$ : Evidence from ‘again’ . . . . .	246
7.4	The Minimal Link Condition and Romance causatives . . . . .	254
7.5	Remaining Problems . . . . .	271
7.5.1	The main verb in English is too high . . . . .	272
7.5.2	Incompatible with Quantifier Float . . . . .	274
7.5.3	PRO, Case Theory and the typology of Infinitives . . . . .	276
<b>8</b>	<b>Rethinking Things</b>	<b>289</b>
8.1	Review . . . . .	289
8.2	Towards Deriving $\bar{X}$ Theory . . . . .	294
8.2.1	Kayne’s “Antisymmetry” hypothesis . . . . .	294
8.2.2	The “Bare Phrase Structure” proposals . . . . .	313
8.3	Embedding the theory in a framework without $\bar{X}$ Skeleta . . . . .	326
	<b>Bibliography</b>	<b>347</b>

# Preface

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

- 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. Chapter 3.

- 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. Chapters 1-5.
- Chomsky, N. "Bare Phrase Structure." In *Government binding theory and the minimalist program*, ed. Webelhuth, G. 383-439. Oxford: Oxford University Press, 1995.

# 1

## The Subject Matter

Linguistic theory, and so syntactic theory, has been very heavily influenced by learnability considerations, 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, physiology or whatever “knowledge” ends up belonging to, 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, for example, as well as the anemic nature of the data available to the child. In addition, the fact that children appear to learn any 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, makes the acquisition of linguistic knowledge look quite different than the acquisition of more familiar domains of knowledge – elementary geometry, for example, or, as you shall see, syntactic theory. How is it that something this complex can, on the basis of such limited information, be acquired with such ease and speed?

### 1.1 Linguistics as learning theory

Chomsky proposed that linguistic theory itself should contribute to solving this puzzle. The classical formulation of his idea (see *Aspects of the Theory of Syntax* and *The Sound Pattern of English*) 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

## 1. The Subject Matter

---

the same structural descriptions to the strings of L that  $G_L$  does. We can think of the learning process as being the selection from the Universe of  $G_{LS}$  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 Gs 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 Gs has structure (i.e. is not so large), and this is the direction that Chomsky takes. This amounts to the claim that there are features of Gs which are built-in: certain properties which distinguish the natural class of Gs from the rest. There is a kind of meta-grammar of the Gs, then, which is sometimes referred to with the label Universal Grammar. Chomsky further hypothesizes that these properties are biologically given: that it is something about the construction of the human brain/mind that is responsible for the fact that the class of Gs are the way they are. This argument, the one that leads from the observation that  $G_{LS}$  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 and given a biological interpretation.

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 imposed by Universal Grammar – which 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 a goal of syntactic theory should be to contribute towards structuring the universe of Gs. He makes some specific proposals about how to envision this in *Aspects of The Theory of Syntax*. He suggests 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, distinguishing the learnable Gs from the unlearnable ones. This criterion can help the analyst 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: select only those models that are ranked highest by the evaluation metric. This model will meet the criterion of explanatory adequacy. It alone will have the properties that enable, under a particular learnability theory, the acquisition of the  $G_L$  that is being described.

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, we might imagine that rules which used fewer symbols could be defined as “simpler” than ones that used a greater number of symbols. Inviolable constraints might be those, for example, expressed as part of  $\bar{X}$  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 a rule construction procedure, and let’s call the companion picture of the acquisition process the “Little Linguist” model.

To take a concrete example, if  $\bar{X}$  Theory – the theory that places limits on phrase structure in Universal Grammar<sup>1</sup> – imposes the constraints expressed in (1) 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.

- (1) a.  $XP \rightarrow \{ (ZP), \bar{X} \}$   
 b.  $\bar{X} \rightarrow \{ \bar{X}, (YP) \}$   
 c.  $\bar{X} \rightarrow \{ X^0, (WP) \}$

(Understand “ $\{\alpha, \beta\}$ ” to signify that  $\alpha$  and  $\beta$  are sisters, “ $(\alpha)$ ” to indicate that  $\alpha$  is optional, and  $\alpha \rightarrow \beta$  to mean that  $\alpha$  immediately dominates  $\beta$ .) 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,  $\bar{X}$  Theory provides information which places an

<sup>1</sup> This will be the subject of the following chapter.

## 1. The Subject Matter

---

absolute cap on the possible phrase markers. In this respect it also illustrates an inviolable constraint.

Let's consider another example, one that Chomsky often points to, involving transformational rules. Transformational rules map one representation to another, typically by way of relocating constituents. Interestingly, it appears that all such rules are "structure dependent." That is, they make reference to the relative structural positions of the moved thing and the position it is moved to. 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 (maybe even provably impossible) to see how such things as  $\bar{X}$  Theory or structure dependence could be learned, they must belong to the features that define the universe of  $G_s$ . And because they are overarching properties of the rules in some  $G_L$ , 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  $G_s$  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:<sup>2</sup>

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))

<sup>2</sup> 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, Humboldt 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 because there is another possible explanation for typological similarities. This is 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. In fact, there is no doubt that the historical relatedness of groups of languages does explain certain typological similarities they bear. It is not implausible to think that the creation of language happened just once in our species, and therefore that all extant languages do have a common parent. Might it be that the universal constraints and typologies that these languages display then merely be the remnants of those, perhaps wholly accidental, properties of that first ancestral language. 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

## 1. The Subject Matter

---

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 a frequently offered proposal is that historical change is a by-product of language acquisition. The idea goes as follows. There are certain “errors” introduced with every generation of language learner, and the result is that the grammar of any one generation is different, perhaps subtly so, from the previous. This infidelity of grammar transmission eventually accumulates changes profound enough to convert one language to another. One might imagine that a variety of external factors might delay or speed up this process. But at root, the engine driving change is 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 evaluation metric. The inviolable constraints will be (among) those things found in all languages. The differences in languages are to be credited to the evaluation 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 Gs 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 some of the details of this emerging picture as it involves syntax.

## 1.2 The evidential basis of syntactic theory

If linguistics is one part of the study of human cognition, in the sense just described, then syntax can be described as that subdiscipline of linguistics which seeks to discover what speakers know about how to arrange the words of their language into meaningful sentences. Because speakers are not conscious of the principles that characterize this knowledge, the syntactician must make recourse to indirect means of determining these principles. The syntactician's first task, then, is to determine how to find evidence that reflects the nature of this knowledge.

One plausible source of relevant information comes from observing how speakers put this knowledge to use. We could, for instance, collect the utterances from some speaker and look for generalizations in these utterances from which evidence about the underlying knowledge-base can be gleaned. This is rarely done, however, as there are few instances of such collections that arise naturally, and to assemble them from scratch is onerous enough to have been avoided. With the exception of studies of prodigious literary figures, there are vanishingly few attempts at linguistic studies that go this route.

More common is to study the linguistic utterances of a group of speakers. This is standardly done by using the dictionary maker's device of combing texts and newspapers for examples. There are several excellent "parsed" corpora of this sort,<sup>3</sup> and even corpora of spoken utterances<sup>4</sup> can be found. With the advent of the World Wide Web, it has become possible to search a very large collection of sentences, and more and more linguists are availing themselves of this resource. This technique has the unique advantage of allowing one to determine frequencies as well. It is possible, for example, to judge how rare some particular arrangement of words is relative to some other, or to find statistically significant correlations between, say, the position of an argument relative to its predicate and the person or number marked on that argument. Some linguistic theories are specifically designed to model these sorts of frequency data.<sup>5</sup>

There are some serious pitfalls to using group corpora, however. One is simply that it obliterates differences among speakers and treats the data as if it were all manufactured by the same grammatical system. Since nothing is known about the producers of these sentences – they may include speakers of different dialects and speakers for whom the language in question is non-native or has been influenced by another language, for instance – this could be a serious source of error. Without

<sup>3</sup> See Marcus, Santorini, and Marcinkiewicz (1993), for example.

<sup>4</sup> See Godfrey, Holliman, and McDaniel (1992).

<sup>5</sup> See the papers in Bod, Hay, and Jannedy (2003) for some recent examples of statistically based corpora studies, and the work of Paul Boersma (e.g., Boersma and Hayes (2001)) for a theory that is designed to model statistical data of this sort.

## 1. The Subject Matter

---

some measure of the heterogeneity of the speakers who produced the corpus, it is very difficult to judge how faithfully it represents the syntactic knowledge of any one of those speakers.

Another shortcoming is that linguistic behavior, even of one individual, is not a faithful projection of the knowledge that that individual has of his or her language. People say sentences whose syntactic form is at odds with what they would otherwise deem well-formed. A significant proportion of any corpus could be made up of such “mistakes,” and indeed it would be prudent to assume so, given the degree to which misshapen sentences populate the utterances of such well-placed contributors to corpora as George W. Bush. There is a distinction between a speaker’s linguistic “performance” and his or her linguistic “competence,” to use the names Chomsky gives to this distinction. Corpora level this distinction.

For these reasons, then, group corpora contain an unknown amount of data that should be weeded out. They contain examples of sentences that are produced by speakers whose grammatical systems differ, and they contain sentences that are not representative of any grammatical system. But group corpora are not only noisy with error, they are also mute about certain kinds of information.

One important piece of evidence that corpora cannot provide concerns where speakers draw the line between impossible and possible forms in their language. This distinction is easiest to elicit in linguistic domains where there are a comparatively small number of relevant forms. For example, the morphological and phonological inventories of any one speaker at any one time is reasonably small and it is therefore salient when a novel morphological or phonological form is introduced. For many such novel forms, speakers are capable of distinguishing those that are admissible members to their languages and those that are not. Most English speakers I have asked, for instance, can tell that *blick* ([blɪk]) is an admissible addition to their lexicon but that *bnick* ([bnɪk]) is not. Presumably this ability to distinguish admissible from inadmissible forms is due to the knowledge speakers have of their language, and so it is an important piece of information about how that knowledge is constituted. A typical way of characterizing this distinction goes as follows. The phonology of a language permits many forms that are not exploited by the lexicon of that language (e.g., [blɪk]). Which of these forms are used and which are not is completely extragrammatical. By contrast, because the phonology of a language limits the forms that are available to that language (e.g., English prevents the onset cluster [bn]) these forms (e.g., [bnɪk] in English) will be blocked from its lexicon. The absence of these forms is determined by the grammar; they are said to be “ungrammatical,” and when they are cited, they are prefixed with the diacritic “\*” to indicate their status.

The same distinction can be elicited for sentences, although because of the larger number of forms involved it is more difficult to recognize a novel sentence.

Consider, by way of illustration, the pair of sentences in (2).

- (2) a. Whenever the earth revolves around its equator, the moon begins to rotate about its axis.
- b. Whenever the earth revolves around its equator, the moon begins itself to rotate about its axis.

I judge (2b) to be an impossible English sentence, and (2a) to be a possible one. Because I read very little science fiction, I think it's likely that both sentences are novel for me, but I do not have the certainty about this that I have about *blick* and *bnick*. I recognize that there are considerably more sentences that I have encountered than there are words I've encountered, and consequently I also recognize that it is likelier that I will mistake a sentence as novel than it is that I will mistake a word as novel. Nonetheless, most linguists would agree that the contrast in (2) is of the same kind that distinguishes *blick* from *bnick*. It does seem unlikely that the distinction could be reduced to one of novelty. After all, I am roughly as certain of the novelty of (2a) as I am of the novelty of (2b) and yet this does not affect the strength of my judgement concerning their Englishness. It seems probable that my ability to judge the difference between (2a) and (2b) traces back to an ability my syntactic knowledge gives me to judge well-formedness.

This distinction between grammatical and ungrammatical forms is important because it seems to tap directly into a speaker's linguistic knowledge. Studying corpora cannot provide what is needed to see this distinction; corpora conflate ungrammatical and grammatical but non-occurring forms. For this reason, and because of its noisiness, I will not use data from corpora in these lectures. But do not forget that corpus studies, and so far as I know only corpus studies, can provide statistical data, for this might be an important resource in forming a complete model.

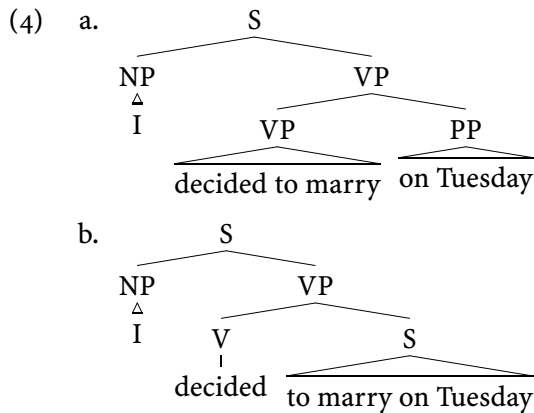
Instead, the central piece of evidence used in these lectures will be elicited grammaticality judgments. This has become the standard tool for linguistic analysis, and much of the literature relies on it. Elicited grammaticality judgments have their own shortcomings. There are special problems attendant with grammaticality judgments of sentences. Because sentences are very complex objects, and are frequently longer than the small memory buffer that our on-line processors are equipped with, there are failures of sentence processing that might easily be mistaken for judgments of ill-formedness. A famous example meant to be illustrative of this distinction comes from strings that are ambiguous with respect to the placement of some late occurring phrase. The pair of sentences in (3) illustrates.

- (3) a. I decided to marry on Tuesday.
- b. I decided that my daughter should marry on Tuesday.

## 1. The Subject Matter

---

Upon reflection, most speakers will recognize that (3a) has two meanings. It can assert that the time of my decision to marry was Tuesday, or it can assert that what my decision was was to marry on Tuesday. As we will see, this ambiguity reflects the fact that (3) maps onto two sentences, whose difference in syntactic structure is responsible for the two meanings. The first meaning corresponds to a structure which groups the words as sketched in (4a), whereas the second interpretation corresponds to the syntactic structure shown in (4b).

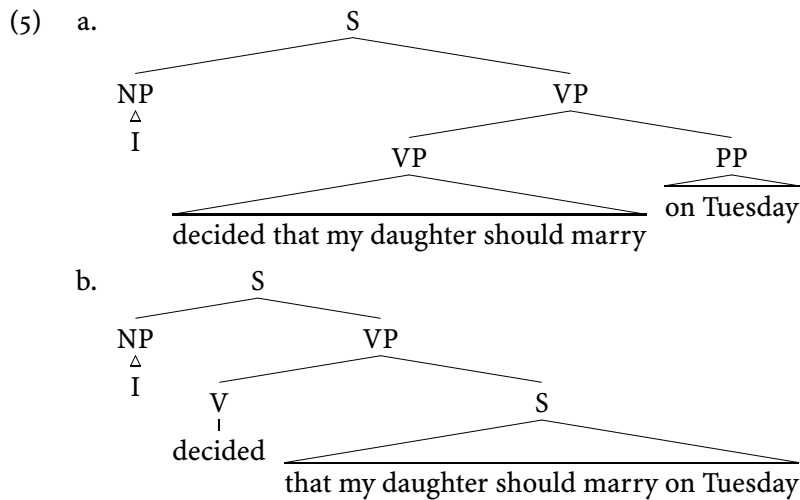


Unlike (3a), (3b) seems to have only the second of these two meanings. It can assert that my decision was for my daughter to marry on Tuesday, but it does not seem to say that the time of my decision was Tuesday. At present, this difference in (3a) and (3b) is thought to be due to constraints of sentence processing, and not the well-formedness conditions of sentences. The relevant difference between these examples is the number of formatives between the word *decided* and the prepositional phrase *on Tuesday*. As that number grows beyond what can be held in working memory, the processor is forced to start making decisions about how to parse the initial portions of the string. These decisions favor a parse in which later material is made part of more deeply embedded phrases. Thus, in the case of (3b) it favors the structure in (5b) over that in (5a) on the facing page. On this account, then, it is not that there is a difference in the syntactic well-formedness conditions which causes speakers' differing judgments about (3a) and (3b). Instead, because of the relative difficulty that (3b) presents to the on-line processor, one of the syntactic representations associated with this string (i.e., (5a)) becomes difficult to perceive. This effect of the on-line processor is what Kimball called "right association."<sup>6</sup>

In general, judgments of well-formedness will not be able to distinguish those sentences that do not conform to the constraints of the grammar from those that

<sup>6</sup> See Kimball (1973), Frazier (1978) and Gibson (1998).





do conform to those constraints but present problems for the on-line processor.<sup>7</sup> There is no simple way of distinguishing these cases; they can be separated only through analysis. In the case of (3), the decision that the effect is not grammatical but, instead, the result of the processor comes partly from finding no good grammatical way of distinguishing the cases and partly from finding that manipulating factors relevant for the processor determine whether the effect materializes.

Another similar difficulty involves the fact that the meanings which sentences convey are typically bound to the context of a larger discourse. Inevitably, then, grammaticality judgments are going to be confounded with whether or not there is a discourse in which that sentence could function. Suppose, for instance, that you are trying to determine the distribution of a process called “VP Ellipsis,” which allows a sentence to go without a normally required verb phrase. VP Ellipsis is responsible for allowing the bracketed sentence in (6) to go without a verb phrase in the position marked “Δ.”

(6) Jerry annoyed everyone that [<sub>S</sub> Sean did Δ ].

If you expose English speakers to the examples of VP Ellipsis in (7), you may find that they judge them ungrammatical.

- (7) a. \* Whomever she did Δ got better.  
 b. \* Everything for her to Δ was hard.

<sup>7</sup> Chomsky and Miller (1963) is an early, and still useful, examination of this distinction.

## 1. The Subject Matter

---

One might be tempted by these examples to the hypothesis that VP Ellipsis is blocked within subjects. But if the examples in (7) are embedded into an appropriate discourse, English speakers will find (7a) well-formed while (7b) remains ungrammatical.

- (8) a. Whomever Sally didn't tutor got worse but whomever she did  $\Delta$  got better.
- b. \* Everything for him to do was easy and everything for her to  $\Delta$  was hard.

The problem with (7a) is that recovering the meaning of the elided VP cannot be done without a larger context, and the grammaticality of sentences with VP Ellipsis in them depends in part on recovering the meaning of the elided VP. There is nothing syntactically ill-formed with the VP Ellipsis in (7a), however, as we see when this context is provided. By contrast, neither the context in (8b) (nor any other that I have found) improves the goodness of (7b). There is something ill-formed about the syntax of this example.

These two problems are similar. In both, the difficulty is in distinguishing judgments of ungrammaticality from other types of ill-formedness. The effect of these difficulties can be lessened if the following two practices are used in eliciting judgments.

First, embed the sentences whose well-formedness you wish to determine in discourse contexts that make the meaning these sentences should have available and salient. This helps remove the second problem.

Second, for every sentence you suspect to be ungrammatical, present your informant with a matching sentence which you suspect to be grammatical. These two sentences – the suspected grammatical and the suspected ungrammatical one – should differ minimally. Your aim should be to remove all differences between these two sentences except for the factor that you suspect is responsible for the ungrammaticality. This will help mitigate processing effects, as the two sentences will end up matched in length and close to matched in complexity. It will also help remove any other confounds which might be responsible for the ungrammaticality of the sentence you wish to test.

These practices are rarely used, unfortunately. The history of syntactic research is littered with dead ends and wrong turns that have resulted from errors in the empirical base as a result. Don't fall victim to these errors. Wherever you can, follow the TWO LAWS OF ELICITATION.

(9) TWO LAWS OF ELICITATION

- a. The sentences for which you elicit a grammaticality judgement should be embedded in a discourse that makes the meaning that sentence would have salient.
- b. Every suspected ungrammatical sentence should be part of a minimal pair, the other member of which is grammatical.

In these lectures, I will sometimes violate (9a) whenever I haven't found a context that improves an ill-formedness judgement. In these cases, my guess is that ungrammaticality of the sentence is either profound or not tied to its information content. Similarly, I will occasionally fail to give a minimal pair when I feel that the ungrammaticality of the sentence involved is dramatic enough to be obvious. For instance, examples such as (10) are so clearly violations of English sentence structure, that I cannot imagine a discourse context that could improve them, nor would minimally contrasting grammatical examples help remove the possibility of a processing effect.

- (10)
- a. \* Many happy the puppies barked.
  - b. \* She talked people to.
  - c. \* He ate should apples.

I do this partly because it will make the exposition cleaner, but obviously also because I am lazy. It would be wise to maintain a healthy skepticism about the data I present with these shortcuts.<sup>8</sup>

There is one last danger in relying on elicited grammaticality judgments, and it is the mundane and familiar one of introducing bias. It is a commonplace among experimental psychologists that eliciting psychological data can involve very subtle ways of introducing bias. Whenever the judgments are less clear than obvious cases like (10), the syntactician should clearly not rely on her or his own judgments. In these cases only judgments elicited from naïve informants will do. And in eliciting those judgments, the syntactician should adopt some of the techniques developed by experimental psychologists. Produce a survey of examples that include the sentences you wish to find judgments for but include irrelevant "fillers" as well. Those sentences should be crafted in accordance with the TWO LAWS OF ELICITATION. Then present this survey to a number of speakers native in the relevant language, controlling as best as possible for dialect variation. Finally, present the items in the survey in a randomized order, mitigating any bias that the order of presentation might introduce. When reporting data, you should also report the number of informants you have used, and make a note of any variation in the judgments you

<sup>8</sup> I will also violate (9) when I am reporting data from the literature in which (9) have not been followed.

## 1. The Subject Matter

---

have encountered. While these safeguards wouldn't satisfy the rigorous numerical criteria of the experimental psychologist, they will go a long way towards removing error and making the data you report comparable to the data someone else gathers.

Grammaticality judgments, then, will be the central evidence used here in uncovering the principles that constitute a speaker's syntactic knowledge. There is one other kind of datum that is important to the syntactician. As we've seen, the syntax of sentences is intimately tied to the meanings they convey. It is the semanticist's job to discover the principles that allow users of language to extract these meanings. One of the central principles of semantics is the law of compositionality.

(11) THE LAW OF COMPOSITIONALITY

The meaning of a string of words,  $\alpha$ , is derivable from the meanings of the strings of words that  $\alpha$  contains.

As a result of the LAW OF COMPOSITIONALITY, there is a regular and productive relationship between the syntactic structure of a sentence and the meaning it conveys. This makes it possible to use the meaning a sentence has to draw conclusions about the syntax of that sentence. This requires the assistance of the semanticist, of course, for an intimate knowledge of the rules of semantic composition are needed to draw these conclusions. In recent years, this source of evidence has grown in importance, and it will be an important component of these lectures.

# 2

## Phrase Structure

Our goal is to model the processes by which arrangements of words are recognized as forming grammatical sentences. As sketched in the previous chapter, this involves discovering how those processes vary from language speaker to language speaker, for only in this way will we get a handle on what features of these processes are universal and on how these processes are permitted to vary. I presuppose that readers of these notes have some familiarity with English, and so we'll begin this task by investigating those processes that are responsible for the English speaker's grammaticality judgments.

Our first observation is that we can get very far in this task using very little information about the words involved themselves. A great deal about the processes that determine well-formed arrangements of words can be characterized using nothing more than the morpho-syntactic "category" that the words belong to. This can be appreciated by virtue of the "Novel Form" argument, which is laid out in (1).<sup>1</sup>

- (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.

<sup>1</sup> I thank Jay Keyser for teaching me this argument.

## 2. Phrase Structure

---

Let's see how far we can get restricting our attention just to morpho-syntactic category.

### 2.1 Substitution Classes

We might start by examining what underlies the notion of morpho-syntactic category, that is the concepts of “noun” and “verb” and so on that are familiar to us from the language arts curriculum. Zellig Harris, exploiting an idea from the Structuralist school of linguistics, argued that morpho-syntactic category should be defined in terms of the linear distribution of words within sentences.<sup>2</sup> Specifically, “noun,” “verb” and so on are “substitution classes” of vocabulary items. They are substitution classes in the sense that there is a set of positions within a sentence into which any member of that class can be substituted preserving the grammaticality of the sentence. For instance, any word that can be grammatically placed in the spot marked with “\_\_\_” in (2) falls within the subset of vocabulary items we know as “nouns.”

(2) the \_\_\_ exists

This is indicated by considering the lists of sentences in (3)-(8).

(3)       The lamp exists.  
          The girl exists.  
          The sky exists.  
          The streetcar exists.

⋮

(4)       \* The happy exists.  
          \* The blue exists.  
          \* The short exists.  
          \* The flat exists.

⋮

(5)       \* The in exists.  
          \* The out exists.  
          \* The from exists.  
          \* The on exists.

⋮

<sup>2</sup> Harris (1946) is an accessible introduction to this procedure.

- 
- (6) \* The swim exists.  
 \* The have exists.  
 \* The ate exists.  
 \* The broke exists.  
 ∴
- (7) \* The slowly exists.  
 \* The apparently exists.  
 \* The always exists.  
 \* The decidedly exists.  
 ∴
- (8) \* The every exists.  
 \* The much exists.  
 \* The no exists.  
 \* The a exists.  
 ∴

As can be seen, this technique picks out a list of words that match what the grammar school curriculum calls nouns, and segregates them from the others. A similar discriminating environment can be devised for each category. For each (major) word class, I've given a distinguishing environment in (9).

- (9) a. have \_\_\_ eaten: ADVERB  
 b. the \_\_\_ thing: ADJECTIVE  
 c. dance \_\_\_ it: PREPOSITION  
 d. in \_\_\_ thing: DETERMINER  
 e. must \_\_\_ there: VERB

Understand (9), and (2) as well, as abbreviating the following claim: there is a sentence that is grammatical which contains “X \_\_\_ Y,” and for which replacing a word of category CATEGORY into “\_\_\_” uniquely preserves grammaticality. So, for instance, (9a) should be understood as claiming that all the ways of completing the sentence (10) involve filling “\_\_\_” with an adverb.

- (10) They have \_\_\_ eaten rutabagas.

On this view, morpho-syntactic categories are simply partitions of the vocabulary into equivalence classes. The labels “noun,” “verb” and so on are merely convenient names for the resulting subsets of vocabulary items.

## 2. Phrase Structure

---

There are a few things about the distinguishing environments in (2) and (9) that should be noted. First, they define substitution classes solely on the basis of adjacent items. We might elevate this to a hypothesis.

- (11) Morpho-syntactic categories can be defined on the basis of what words they can be adjacent to.

Second, the environments in (9) partition the vocabulary in ways that your language arts curriculum may not have. For instance, the Determiner class picked out by (9d) does not include *much* or *many*. There aren't grammatical sentences that contain *in much thing* or *in many thing* as a substring. One reaction to this would be to allow *much* and *many* to belong to different word classes than *every*, *the*, *a*, and so on. We could admit the two additional word classes,  $DET_{mass}$  and  $DET_{plural}$ , defined over the environments in (15).

- (12) a. in \_\_\_ stuff:  $DET_{mass}$   
b. in \_\_\_ things:  $DET_{plural}$

This is a straightforward application of the procedure for defining morpho-syntactic category that Harris's program offers, and it is one direction that syntactic theorists go.

There is another reaction to these data, however, and it is the one I shall follow. It's clear by comparing the environments that define Determiner and  $DET_{plural}$  that what distinguishes these two word classes is whether the word that follows is plural or singular. The difference between singular and plural is a semantic one, and so we should tie the difference between Determiners and  $DET_{plural}$  eventually to a semantic primitive. It is also a semantic difference, although a less familiar one, that distinguishes the Determiner and  $DET_{mass}$  categories. Words such as *stuff* refer to entities which do not contain clearly delineated parts; whereas words like *things* refer to entities that do. If one collects one thing and adds it to another thing, then the result is an entity named after its two parts: *things*. The same is not true of two collections of stuff. Words that refer to entities that can be piled together, or taken apart, in the way that things can are called "count nouns," while those that cannot are called "mass nouns." The difference between the Determiner and  $DET_{mass}$  classes is just whether the term that follows them is mass or count. There is a clearly semantic generalization to be captured in distinguishing these classes of determiners, and we should strive to capture these generalizations in our grammar.

There are generalizations hidden in the environments in (2) and (9) as well, but it is not at all clear that these are semantic generalizations. To see this generalization, consider the following series of distinguishing environments for the word class noun that are very similar to (2).



- (13) a. the \_\_\_ eats  
b. some \_\_\_ knows  
c. a \_\_\_ exists  
d. few \_\_\_ is  
e. every \_\_\_ ate  
f. no \_\_\_ exists  
g. some \_\_\_ has  
h. every \_\_\_ put  
i. a \_\_\_ screamed  
j. few \_\_\_ drove  
k. and so on

The generalization in this list is that the words flanking the environment in which nouns are restricted are themselves of a word class; each member of this list fits the schema in (14).

- (14) DETERMINER \_\_\_ VERB

Each of the environments in (9) can be similarly converted into a generalization that makes reference to morpho-syntactic category.

- (15) a. VERB \_\_\_ VERB: ADVERB  
b. DETERMINER \_\_\_ NOUN: ADJECTIVE  
c. VERB \_\_\_ NOUN: PREPOSITION  
d. PREPOSITION \_\_\_ NOUN: DETERMINER  
e. INFL \_\_\_ PREPOSITION: VERB

(NB: The word *must* belongs to a morpho-syntactic category with a small set of members; I've labeled it INFL in (15e). We'll soon encounter the evidence for this category.) At present it is not possible to reduce this generalization to a semantic one. That is, there is no known method of defining morpho-syntactic categories in semantic terms.<sup>3</sup> At present, the best that can be done is to define morpho-syntactic categories in the terms that Zellig Harris gave us: substitution classes. The generalizations underlying (2) and (9) are at present irreducibly syntactic, then.

Notice that converting (9) to (15) claims that the particular lexical items chosen will not matter. But, as we've just seen, it does matter: whether the noun in (15d) is count or mass or singular or plural will determine which of the Determiner,

<sup>3</sup> For a recent attempt to define some of the major categories in terms that verge on semantic, see Baker (2003).

## 2. Phrase Structure

---

$Det_{mass}$  and  $Det_{plural}$  classes are well-formed in this position. To make this step, then, requires factoring out the semantic information that controls these alternations.

One reaction to the differences among Determiner,  $Det_{mass}$  and  $Det_{plural}$ , then, is to segregate the kinds of information that together determine the distribution of words. This is the path we shall take. We assign to the semanticist the task of explaining the wholly semantic part of this job: why, for instance, *much* can be left-adjacent to a mass noun but not a count noun. In general, it is not trivial to know when it is the semantics or the syntax that is responsible for coöccurrence restrictions like those in (2) and (9), and the line is constantly being questioned. Harris, it seems, believed that virtually none of it was semantic, whereas present-day categorial grammarians push in the direction of removing the syntactic contribution. I'll chart a course that is somewhere in the middle.

Morpho-syntactic categories, then, are defined in syntactic terms. They are subsets of the vocabulary that can be substituted for each other in particular positions within a grammatical sentence preserving grammaticality. Moreover, the particular positions can be characterized in terms of adjacent morpho-syntactic categories. The first step in characterizing the grammaticality judgments of some speaker is recognizing that the vocabulary of that speaker is partitioned in this way.

### 2.2 Phrases

Certain strings of categories also have a distribution within sentences that can be defined in terms of adjacent items. For example, the string  $D(\text{eterminer})+Adj(\text{ective})+N(\text{oun})$  can appear immediately after a preposition and immediately preceding the 's which marks the "possessive."

- (16) \_\_\_ 's & P \_\_\_ :  $D+Adj+N$
- a. I talked to **the happy woman**.
  - b. **the happy woman's** friend

This string can also be "coördinated" with another identical string of categories. Coördination involves the use of words called "conjuncts," words such as *and*, *or*, *nor*, etc. Thus, we find examples like (17) but not (18).

- (17) the happy woman and an unhappy man
- (18) a. \* the angry and an unhappy man  
b. \* the and an unhappy man

Finally, with respect to all these distributional tests, the strings  $D+N+P(\text{reposition})+N$ ,  $N+P+N$ ,  $Adj+N$ ,  $N$ , 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. That is, this family of strings is a substitution class in the same sense that morpho-syntactic categories are.

This family of strings is called a “phrase,” and we can write a Phrase Structure Rule to describe which strings belong to this family. In the case at hand, this rule might look like:

$$(19) \quad \alpha P \rightarrow (D) (\text{Adj}) N$$

Understand material enclosed within “( )” to be optional; (19) therefore generates the set of strings: D+Adj+N, D+N, Adj+N and N.

This leaves out the strings D+N+P+N and N+P+N; but these strings, involve another phrase, made up of the string P+N. This string, along with any string that conforms to the template P or P+ $\alpha$ P or P+P or P+P+ $\alpha$ P has the defining distribution in (20).

- (20) A \_\_\_ & P \_\_\_
- a. I stood **around**.
  - b. I knew the man **by Mary**.
  - c. I remain disliked **by Mary**.
  - d. I stood next **to Mary**.

Like  $\alpha$ Ps,  $\beta$ Ps may be coordinated with other  $\beta$ Ps, but not with other  $\alpha$ Ps, as the following examples illustrate.

- (21) 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  $\alpha$ Ps, this family of strings constitutes a substitution class.

Putting this together, we come up with the Phrase Structure rules in (22).

- (22) a.  $\alpha P \rightarrow (D) (\text{Adj}) N (\beta P)$
- b.  $\beta P \rightarrow P (\alpha P)$
- c.  $\beta P \rightarrow P (\beta P)$

It is customary to collapse the two rules in (22b,c) to (23).

$$(23) \quad \beta P \rightarrow P \left\{ \begin{array}{l} (\alpha P) \\ (\beta P) \end{array} \right\}$$

## 2. Phrase Structure

---

Note that (22a) and (23) together have the property of being 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 phrase structure 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 are a substitution class. The environment that defines them is (24).

- (24) Infl —
- a. I should **eat rutabagas**.
  - b. I will **talk to Mary**.
  - c. I will **tell Mary about rutabagas**.

“Infl” is a morpho-syntactic category that includes *should*, *will*, *must*, *would*, *will*, *can*, *could* and a few other words. Like  $\alpha P$ s and  $\beta P$ s, coördination treats members of this family as alike and distinct from  $\alpha P$ s and  $\beta P$ s.

- (25) 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:

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

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  $\gamma P$ , a noun the only requisite member of an  $\alpha P$  and a preposition is all that's required to make a  $\beta P$ . Further, the converse also turns out to be true: whenever there is a preposition, there is a  $\beta P$ , wherever a noun is found, there is an NP, as so on. Thus, nouns and  $\alpha P$ , prepositions and  $\beta P$ , verbs and  $\gamma P$  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 to always hold. 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  $\alpha P$ , NP,  $\beta P$  PP and  $\gamma P$  VP. Thus, we now have the rules in (27).

- (27) a. NP → (Det) (Adj) N (PP)  
 b. PP → P  $\left\{ \begin{array}{l} \text{(NP)} \\ \text{(PP)} \end{array} \right\}$   
 c. VP → V (NP) (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:

- (28) AP → A (PP)

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

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

The coördination test also treats A and A+PP strings as being the same:

- (30) 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.

- (31) S → NP Infl VP

The morpho-syntactic category that sentences are in a one-to-one relation with is Infl,<sup>4</sup> and so in keeping with the convention of naming phrases after their heads, we should change (31) to (32)

- (32) IP → NP Infl VP

With this rule we have finally come to the task of characterizing the grammaticality judgments of English speakers. For any speaker of English whose vocabulary has been partitioned into noun, verb, preposition, adjective, determiner and Infl,

<sup>4</sup> In this case, however, unlike what we found for the VP, NP and AP rules, Infl is not the only obligatory member of a sentence. It is presently controversial whether sentences are the only phrases that have this property. We will see in later lectures that there are syntactic reasons for the obligatory presence of NP, and likely semantic ones for the obligatory presence of VP.

## 2. Phrase Structure

---

(32), with the rules in (27), characterizes those strings of words that will be judged grammatical.

This is just a first step, of course. We have hundreds of pages left. In fact, it's possible to see something wrong with (32) right away. It says that no sentence can fail to have an Infl between NP and VP, but if Infl are just words such as *can*, *could*, *will*, and so on this is obviously wrong. There are grammatical sentences aplenty that fail to have these words in them; (33) for instance.

(33) Jerry walked.

Where is the Infl between *Jerry* and *walked* in this sentence?

If we look hard, we find that sentences are, in fact, in a one-to-one correlation with a category, but that that category includes not just words, but bound morphemes as well. Consider the sentences in (34).

- (34)
- 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 **leaving**].

The boldfaced terms have similar distributions: they are found either immediately preceding the verb (if they are free) or affixed onto the following verb (if they are bound). Every sentence has one of these, and so these terms meet the criteria of being the head of a sentence. To explain how it is that those Infs which are bound morphemes materialize affixed onto the following verb, we will have to invoke a process that goes beyond phrase structure rules. Let us put off doing this.

As we gather more detail about the shapes of grammatical English sentences, we will need to make quite a number of additions to these rules. In fact, to be exhaustive about this proves to be a task beyond what we can manage here; we should consider this an open-ended process. Nonetheless, I want to gather a little more detail than we now have.

I'll begin by adding a couple of phrases to our inventory. One of these is a sort of "sentence" found in examples like (35).

- (35)
- 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 IPs. The word *that* is called a “complementizer” and it is the head of the phrase found in these sentences. This phrase, or clause as sentence-like phrases are often called, is a “Complementizer Phrase” (CP). CPs conform to the requirements of the following Phrase Structure rule.

(36)  $CP \rightarrow C IP$

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

- (37) 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. This yields the following battery of rules.

- (38) a.  $IP \rightarrow \left\{ \begin{array}{l} NP \\ CP \end{array} \right\} I VP$   
b.  $NP \rightarrow (D) (AP) N (PP) (CP)$   
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. Note too that I’ve brought the NP rule into conformity with the principle of endocentricity. Our earlier rule –  $NP \rightarrow (D) (Adj) N (PP)$  – permitted an adjective without an adjective phrase. I’ve replaced “(Adj)” with “(AP)” to correct for this. We’ll see the empirical support for that change below.

The second phrase we’ll need are ones headed by adverbs. Adverbs are a word class that is sometimes defined on the position left adjacent to adjectives, as in (39).

- (39) a. A **very** happy child  
b. The **extremely** large boat

They can also be found left adjacent to verbs, as in (40).

- (40) a. I have **deliberately** misled.  
b. I have **noticeably** erred.

Interestingly, to a large extent, the set of adverbs that can be immediately preceding the verb is a proper subset of those that can immediately precede the adjective.

- (41) a. A **deliberately** angry child  
b. The **noticeably** large boat

## 2. Phrase Structure

---

- (42) a. \* I have **very** misled.  
b. \* I have **extremely** erred.

Here is one of those places where we must judge whether we have two morpho-syntactic categories or we have a semantic problem. To judge from these few examples, it seems that the adverbs which indicate an “extent” or “measure” are fine in the pre-adjectival position but not in the preverbal position. This description of the contrast makes reference to the meanings of the words involved, and so it could reflect a semantic fact. On the other hand, it may be that we **are** looking at a contrast due to morpho-syntactic category, but one that correlates with (or perhaps is defined by?) these semantic factors. I don’t know how to decide between these alternatives. I will gamble that this is a category distinction like the others we are examining. Let’s call the category of words that I’ve characterized as denoting an extent, “degree,” and reserve the term “adverb” for the others.

The class of Degree words can also show up in combination with adverbs, as illustrated by (43).

- (43) a. Mary **very** quickly walked through the full classroom.  
b. Mary **extremely** loudly declared her interest in Dutch phonemics.

Indeed, wherever an Adverb can be found, so also can a Deg+Adverb string. Our phrase structure rule for adverb phrases will be (44), therefore.

- (44) AdvP → (Deg) Adv

If we admit the category Degree, then the principle of endocentricity tells us that there must be a phrase that has the same distribution. It is very difficult to identify a (non-trivial) string that has the same distribution as Degree words. Those strings are most easily seen in so-called comparative constructions, a complex phenomenon that we will hopefully manage to skirt in these lectures. As a stop-gap, let us do with the questionable rule in (45).

- (45) DegP → Deg

We should modify (44) accordingly.

- (46) AdvP → (DegP) Adv

The existence of degree phrases and adverb phrases now requires that we modify some of our other rules so that they can be positioned within sentences correctly. Adjective phrases will have to be changed to allow for degree phrases within them, as in (47).

- (47) AP →  $\left\{ \begin{array}{l} \text{(DetP)} \\ \text{(AdvP)} \end{array} \right\} A \text{ (PP) (CP)}$



And refinements to the rule that characterizes verb phrases will be necessary to position adverb phrases within them.

(48)  $VP \rightarrow (AdvP) V (NP) (PP) (CP)$

There are some other changes to the VP rule that are necessary. Note, for instance, that VPs may occur immediately following a verb, as in (49).

- (49) 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. For instance, *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:

(50)  $VP \rightarrow (AdvP) V \left\{ \begin{array}{l} (NP) \quad (PP) \quad (CP) \\ \quad \quad \quad VP \end{array} \right\}$

Further, it is possible to find APs embedded within VPs; some examples are:

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

When APs follow verbs, they may be preceded by, at most, a PP, as in (52).

(52) Jerry seem [<sub>PP</sub> to Bill] [<sub>AP</sub> happy with his rutabagas].

So we change the rule that characterizes VPs to:

(53)  $VP \rightarrow (AdvP) V \left\{ \begin{array}{l} (NP) \quad (PP) \quad (CP) \\ \quad \quad \quad VP \\ \quad \quad \quad (PP) AP \end{array} \right\}$

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 (54).

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

## 2. Phrase Structure

---

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 within NPs. We will revisit this idea, but for now let's imagine that determiner phrases are made up of nothing but determiners.

(55)  $DP \rightarrow Det$

We'll therefore need to update the phrase structure rule that forms NPs. But before we do this, let's consider strings like those in (56).

- (56) 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, as (57) illustrates.

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

One very typical explanation for this is to understand determiners and possessives as competing for the same position. 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 the curly braces abbreviation in (58).

(58)  $NP \rightarrow \left\{ \begin{array}{l} (DP) \\ (NP's) \end{array} \right\} (AP) N (PP) (CP)$

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.

(59)  $\alpha \rightarrow \alpha \text{ Conj } \alpha$

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*, *or* and *but*.

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

(60) a.  $IP \rightarrow \left\{ \begin{array}{l} NP \\ CP \end{array} \right\} I VP$

- b.  $NP \rightarrow \left\{ \begin{array}{l} (DP) \\ (NP's) \end{array} \right\} (AP) N (PP) (CP)$
- c.  $VP \rightarrow (AdvP) V \left\{ \begin{array}{l} (NP) \quad (PP) \quad (CP) \\ \quad \quad VP \\ \quad \quad (PP) AP \end{array} \right\}$
- d.  $DP \rightarrow Det$
- e.  $DegP \rightarrow Deg$
- f.  $AdvP \rightarrow (DegP) Adv$
- g.  $AP \rightarrow \left\{ \begin{array}{l} (DegP) \\ (AdvP) \end{array} \right\} A (PP) (CP)$
- h.  $CP \rightarrow C IP$
- i.  $PP \rightarrow P \left\{ \begin{array}{l} (NP) \\ (PP) \end{array} \right\}$
- j.  $\alpha \rightarrow \alpha Conj \alpha$

An interesting property of the phrases defined in (60), and one which Harris noted, is that none of them phrase itself with its head. Harris argues that this is a desirable property of these rules. 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 (61) and not (62).

$$(61) \quad NP \rightarrow (DP) (AP) N (PP)$$

$$(62) \quad N \rightarrow (DP) (AP) N (PP)$$

Or, to put it somewhat differently, we do not want these rules to be recursive with respect to their head. The phrases we've encountered so far all have this property, but not all phrases do. We turn to these others now.

### 2.3 $\bar{X}$ phrases

There are substitution classes that pick out strings which are recursive on themselves. That is: these phrases are headed by themselves. They are little phrases inside those we've identified so far. For example, in the position marked by “\_\_\_” in (63), we find the family of strings in (64). Some examples are in (65).

## 2. Phrase Structure

---

- (63) Det \_\_\_ V
- (64) {N, AP N, N AP, N PP, AP N PP, N PP AP, AP AP N, N PP PP, AP AP N PP, AP AP N PP PP, ...}
- (65) the woman left.  
 the happy woman left.  
 the woman unhappy with the lecture left.  
 the happy woman with a hat left.  
 the woman with a hat unhappy with the lecture left.  
 ⋮

And coordination also reveals that this set of strings forms a family:

- (66) The woman and happy man left.  
 The happy woman and man with a hat left.  
 ⋮

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:

- (67) a. The woman left.  
 b. \* 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. This family is recursive with respect to a substring that includes its head. So, we set up something like this:

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

Note how these rules encode the "optionality" of AP and PP differently than the optionality of DP. 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. We will return to it in just a moment.)

We find the existence of very similar subphrases within VPs as well. Consider, for instance, the environment in (69), which permits the family of strings in (70), as (71) exemplifies.

- (69) NP \_\_\_ CP
- (70) {V, AdvP V, V AdvP, V PP, AdvP V PP, V PP AdvP, V AdvP PP, AdvP V PP PP, AdvP AdvP V, AdvP AdvP V PP, V PP PP, ...}
- (71) Sally said that Jerry left.  
 Sally quickly said that Jerry left.  
 Sally quickly said to Peter that Jerry left.  
 Sally said to Peter quickly that Jerry left.  
 Sally said quickly to Peter that Jerry left.  
 Sally carefully said to Peter on Tuesday that Jerry left.  
 ⋮

And, as before, coördination recognizes this family.

- (72) 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:

- (73) a.  $VP \rightarrow \bar{V}$   
 b.  $\bar{V} \rightarrow AdvP \bar{V}$   
 c.  $\bar{V} \rightarrow \bar{V} AdvP$   
 d.  $\bar{V} \rightarrow \bar{V} PP$   
 e.  $\bar{V} \rightarrow V$

These rules leave out the expansions of VP which introduce NPs, CPs, APs, and VPs. Moreover, the first of these rules says that VPs are  $\bar{V}$ s and nothing more, which raises the obvious question why we posit  $\bar{V}$ s here at all. We would get the same result by dispensing with the first of these rules, and replacing  $\bar{V}$  with VP throughout the remainder. We will soon see, however, that in certain situations there is a term that can show up which appears to be dominated by VP but not  $\bar{V}$ . I'll keep these rules in anticipation of that situation.

A similar situation arises in Adjective Phrases too. If we examine the environment in (74) we discover that it characterizes the set of strings in (75).

- (74) V \_\_\_ CP
- (75) {A, Deg A, Deg Deg A, A PP, Deg A PP, Deg A PP PP, ...}
- (76) Sean is **happy** that syntax is cool.  
 Sean was **happy on Tuesday** that syntax is cool.  
 Sean was **very happy on Tuesday in this class** that syntax is cool.  
 ⋮

## 2. Phrase Structure

---

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

- (77) A child **happy with her guardian** and **well-rested** is unlikely to cause trouble.  
A child **happy with her guardian** and **completely well-rested** is unlikely to cause trouble.  
A child **thoroughly unhappy in a zoo** and **angry at her guardian** is likely to cause trouble.

⋮

We need to revise the AP rule to something like:

- (78) a.  $AP \rightarrow \bar{A}$   
b.  $\bar{A} \rightarrow \text{Deg } \bar{A}$   
c.  $\bar{A} \rightarrow \bar{A} PP$   
d.  $\bar{A} \rightarrow A$

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

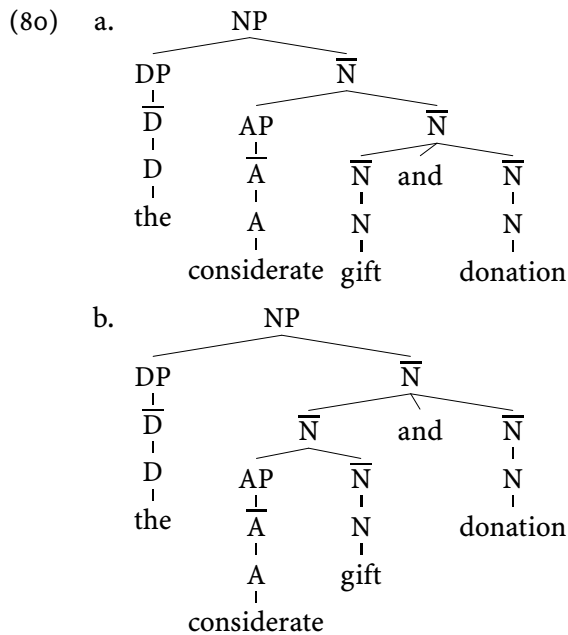
There is a feature of this method of representing these subfamilies that I would like to draw attention to now. It allows for two separate parses of examples such as (79).

- (79) the considerate gift and donation

It is possible to produce this string either by grouping *considerate* and *gift* into one  $\bar{N}$  and conjoining that with an  $\bar{N}$  consisting of just *donation*, or it is possible to conjoin *gift* and *donation* into one  $\bar{N}$  and then group that phrase with *considerate* into an  $\bar{N}$ . It is easy to represent these two parses by way of "phrase marker trees," which graphically elucidate the constituent structure of strings. The two ways of producing (79) are represented by the trees in (80) on the next page. 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. An loose paraphrase of these two meanings might be as given in (81).

- (81) a. the things which are considerate and which are, first, a gift and, second, a donation  
b. the things which are, first, a considerate gift and, second, a donation

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. For instance, if we add one more adjective to the left of the coördinated nouns, as in (82), our rules allow for a total of three parses (shown in (83) on the following page) and there are three meanings as well (as indicated in (84)).

(82) the considerate big gift and donation

- (84) a. the things which are considerate and big and are also a gift and a donation.  
 b. the things which are considerate and are also a big gift and a donation  
 c. the things which are a considerate big gift and a donation

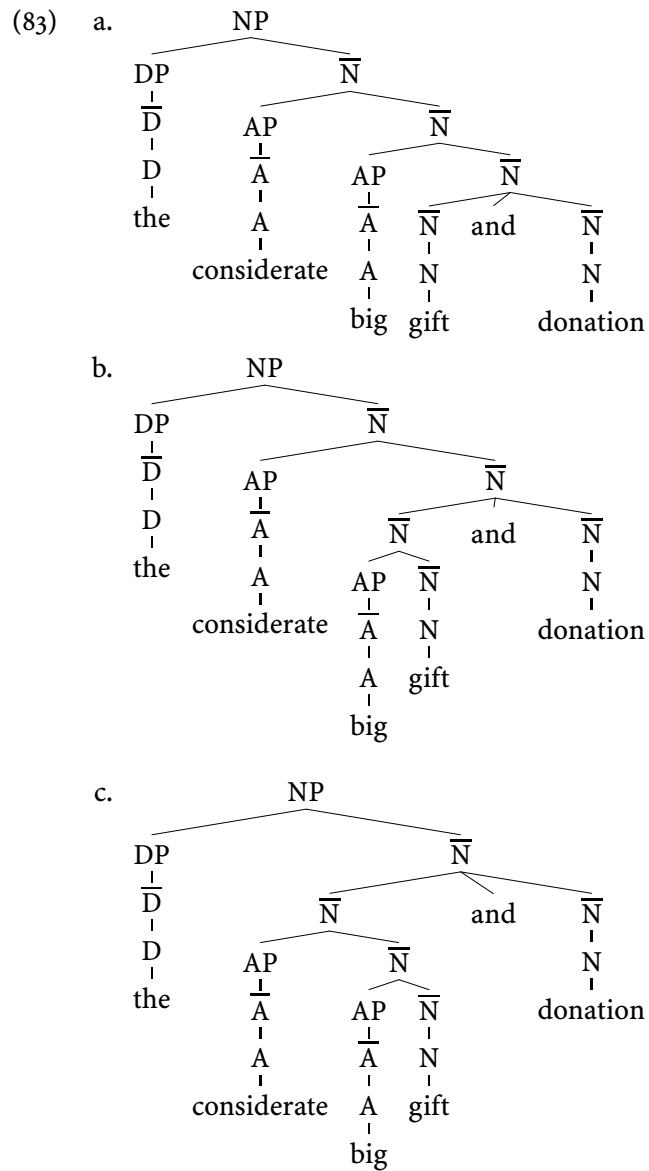
Furthermore, the meanings vary in a predictable way with the linear order that these terms are arranged in. Thus, for instance, putting the second adjective to the left of the coördinated nouns creates the three meanings listed in (84), whereas putting the second adjective to the left of just the rightmost noun, as in (85), produces just two readings and they are (86).

(85) the considerate gift and big donation

- (86) a. the things which are considerate and both a gift and a big donation  
 b. the things which are a considerate gift and a big donation

## 2. Phrase Structure

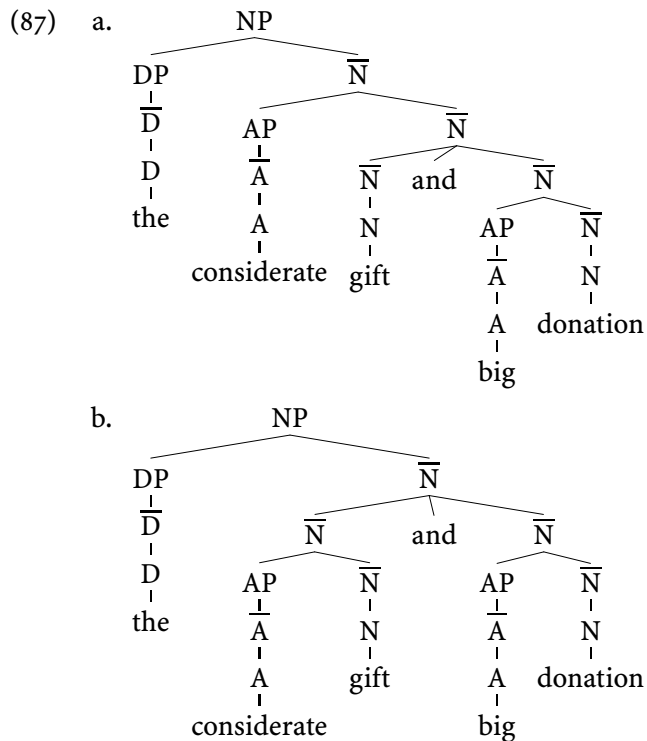
---



This is predictable in the sense that our characterization of these strings would deliver just the two parses for (85) shown in (87) on the next page.

This correspondence should give us some courage that we are on the right track in characterizing the infinite strings under discussion in terms of recursive phrases. It provides a set of structures that are in correspondence with what look like a par-





allel set of meanings. Our next step should be to flesh out this correspondence, but we have some work still to do in characterizing these basic facts about grammaticality judgments. So let's return to that task.

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:

- (88) 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.

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

## 2. Phrase Structure

---

So, let's convert the AdvP rule to (90).

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

Like the AP and VP rules, this battery of rules equates AdvP with  $\overline{\text{Adv}}$  and so makes mysterious why they are called different things.

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 (91).

- (91) 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:<sup>5</sup>

- (92) Jerry evidently won't speak.  
Jerry evidently deliberately won't speak.  
Jerry evidently won't speak deliberately.  
Jerry evidently occasionally deliberately won't speak.  
Jerry evidently won't speak occasionally deliberately.

(These are all somewhat strained, I grant you, but I think still grammatical.) This calls for a change along the lines in (93).

- (93) a.  $\text{IP} \rightarrow \left\{ \begin{array}{l} \text{NP} \\ \text{CP} \end{array} \right\} \bar{\text{I}}$   
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 battery of rules, unlike the others we've formulated, the  $\bar{\text{X}}$  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:

<sup>5</sup> But haven't we already characterized strings like the third and the fifth in (92) as adverbs introduced by a recursive  $\bar{\text{V}}$  rule? Do we really need to also let sentence-final adverbs be introduced by a recursive  $\bar{\text{I}}$  rule? The answer typically given is: yes. But you might want to decide for yourself what the answer to this should be.

(94)  $\bar{V} \rightarrow V \text{ (NP) (CP)}$

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:

(95)  $\bar{A} \rightarrow A \text{ (CP)}$   
 $\bar{N} \rightarrow N \text{ (CP)}$

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 “ $\bar{X} \rightarrow X \dots$ ” (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.<sup>6</sup>

For example, the term *one* prosodically looks like a word, but semantically derives its meaning by being anaphoric to an  $\bar{N}$ .

- (96) 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  $\bar{N}$ , but is also syntactically an  $\bar{N}$ , is because of the contrast in (97), a contrast which also supports our treatment of the non-recursive parts of NP.

- (97) a. I will examine the long proof that language exists if you will examine the short one.  
*one* = “proof that language exists”

<sup>6</sup> “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*.

## 2. Phrase Structure

---

- b. \* I will examine the long proof that language exists if you will examine the one that it doesn't.  
*one* = “long proof”

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 (98).

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

The difference between (98) and (97b) is just whether the material that combines with *one* is allowed to be a sister to an  $\bar{N}$  or not: PPs are (look at (96c)), and CPs aren't.<sup>7</sup>

Similarly, the  $\bar{V}$  family can be anaphorically connected to other  $\bar{V}$ s, but in this case the phonological manifestation of the anaphor is silence, which will be designated with “ $\Delta$ ” in what follows.

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

- (100) a. \* Although Sally shouldn't  $\Delta$  Chicago, Jerry must leave New York.  
 $\Delta$  = “leave”  
b. \* Although Sally didn't  $\Delta$  that she was tired, Jerry said that he would sleep.  
 $\Delta$  = “say”

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

<sup>7</sup> We will see, in just a moment, that this does not turn out to be a categorial distinction, however – so be forewarned. In particular, it will emerge that semantic function of the PP or CP determines how it behaves with regard to this test, and not the mere fact that the phrase is a PP or CP.

- (101) 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  $\Delta$  about George, Jerry will carelessly talk about Sal.  
 $\Delta$  = “carelessly talk”

So we should change these rules to:

- (102)  $\bar{N} \rightarrow N$  (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 seems true sometimes, as in (103) and (104).

- (103) a. Jill ate it at noon.  
 b. \* Jill ate noon it.  
 (104) a. Jill ate spätzle at noon.  
 b. \* Jill ate at noon spätzle.

But for other cases it seems uncertain, or downright wrong, as in (105) and (106).

- (105) a. Jill ate the rotting kumquats.  
 b. Jill ate at noon the rotting kumquats.  
 (106) 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.

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 DP are too anemic for us to see that structure, so we don’t know, empirically, whether or not they have it.<sup>8</sup> 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.

<sup>8</sup> We will eventually see that PP and DP do, but it requires more exotic constructions than we are now prepared for.

## 2. Phrase Structure

---

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.

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

These all conform to the following shape.

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

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

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,  $\overline{V}$  can have an AP adjoined to it, as in (109).

(109) Sandy saw a man today [angry at her].

And, as noted earlier, an  $\overline{N}$  can have certain kinds of CPs adjoined to them, see (110).

(110) 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.

## 2.4 Arguments and Modifiers

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:

- (111) a. Jerry danced with pickles.  
b. Jerry danced at noon at midnight.  
c. Jerry slowly stood still.  
d. a green idea

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 (112).

- (112) 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 the material that follows them incorrectly.

- (113) 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 (114).

## 2. Phrase Structure

---

- (114) 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 events denoted by "I left" refers 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 (115).

- (115) 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 is considerably more constant.

There's a way of talking about argumenthood that is commonplace, and which we inherit from Gruber (1965). Gruber, and Fillmore (1968) in a similar paper, 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.'

- (116) Sandy sent his book to Sean.  
*Source Theme Goal*

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

- (117) Sandy showed his book to Sean.  
Sandy pleases Sean.

It should be noted, however, that there are cases which even metaphorical extensions of the logic of motion look unlikely to characterize; (118) for example.



- (118) Sandy made Sean happy.  
Sandy kissed Sean.  
Sandy wrote his book for Sean.

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  $\theta$ -role.

The relation between verbs and their arguments expressed by  $\theta$ -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.  $\theta$ -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  $\theta$ -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.

- (119) a. Martha denied that John has left.  
b. Martha said that John has left.  
c. \* Martha wonders that John has left.
- (120) a. \* Martha denied 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 functions that  $\theta$ -roles typically name, I will use the language of  $\theta$ -roles to describe this relation too. I am keeping here with the sloppy usage often employed in the literature.

What the contrast between (112) and (113) indicates is that verbs also specify what "category" their argument must be. So, as the pairs below shows, this is another property which distinguishes verbs.

- (121) a. Jerry likes Mary.  
b. \* Jerry likes to Mary.

## 2. Phrase Structure

---

- (122) 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.<sup>9</sup>

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:

- (123) a. Although Sally didn't  $\Delta$  Tuesday, she will dance Monday.  
b. \* Although Sally didn't  $\Delta$  Tuesday, she will remember Monday.  
(124) a. Although Sally won't  $\Delta$  on the bus, she will kiss her in the car.  
b. \* Although Sally won't  $\Delta$  on the bus, she will put her in the car.

Because the phrase following the verb is an argument in (123b) and (124b), it must be within the  $\bar{V}$  which elides, whereas in (123a) and (124c), 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.

- (125) 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.  
(126) 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 (125). 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 (126) 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 (125) and (126) is that the phrases following *do so* in (126) are arguments, whereas those in (125) aren't.

<sup>9</sup> The 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  $\theta$ -roles we just reviewed can be identified in NPs, Jackendoff's claim seems correct here as well.<sup>10</sup>

- (127) a. I'll listen to your long, careful discussion of it, if you'll listen to my short *one*.  
*one* = "careful discussion of it"
- b. \* I'll listen to your long, careful discussion of it, if you'll listen to my short *one* of it.  
*one* = "careful discussion"
- c. I'll listen to your long, careful discussion in class, if you'll listen to my short *one* in the office.  
*one* = "careful discussion"

The contrast between (127a) and (127b) will follow if *of it* must be positioned within the smallest  $\bar{N}$ . The contrast between (127b) and (127c) corresponds to the differing argument-status of the PPs involved: *of it* is more strongly perceived as an argument of *discussion* than is *in class*. As with the *do so* and  $\bar{V}$  Ellipsis facts, then, this contrast supports the hypothesis that arguments and modifiers are fit into phrases in different positions.<sup>11</sup>

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  $\theta$ -role, and allow  $\theta$ -roles to be assigned only to complement positions. We need also to describe the fact that when a verb has a  $\theta$ -role, there must be an argument present in the syntax which bears that  $\theta$ -role. It is customary to divide this task into two parts, which can be expressed as follows:

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

<sup>10</sup> There is evidence that nouns assign  $\theta$ -roles only when they have a verb-like use; that is, when they are used to describe processes or events, and not when they are used to name things (see Grimshaw (1990)). The examples in (127) are constructed with this in mind.

<sup>11</sup> Baker (1978) is, perhaps, the first to argue from *one* anaphora for this conclusion about where arguments are positioned within NPs.

## 2. Phrase Structure

---

It is usual to strengthen the Theta Criterion to a bijection, because of cases like (129).

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

Without constraining the THETA CRITERION to a bijection, we might expect (129) to get such an interpretation since presumably the NP *John* could name the object which serves as argument of both the Theme and Goal  $\theta$ -role. So we change this to (130).<sup>12</sup>

(130) THE THETA CRITERION

- a. For every  $\theta$ -role there is exactly one position to which that  $\theta$ -role is assigned.
- b. For every  $\theta$ -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 (131).

(131) For every  $\theta$ -role assigned by some X, X c-selects the phrase that bears that  $\theta$ -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.

- (132) 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  $\theta$ -role, or, alternatively, that something relaxes the condition which forces every  $\theta$ -role to be assigned to a position holding an argument. 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  $\theta$ -roles for certain predicates to not be assigned to a position.

<sup>12</sup> The THETA CRITERION is also often formulated in terms of a bijective relation between  $\theta$ -roles, or  $\theta$ -positions, and arguments. That is, it is sometimes written to say: "For every  $\theta$ -role (or positions) there is exactly one argument and for every argument there is exactly one  $\theta$ -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  $\theta$ -role, whereas the alternative does. I've decided to place this requirement in another principle, which we'll come to shortly.

Some have suggested that (128b) and (131) should be collapsed, and in particular, that there is a means by which the categorial type of some argument can be determined from its  $\theta$ -role.<sup>13</sup> 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  $\theta$ -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  $\theta$ -roles. These  $\theta$ -roles can sometimes be borne by NPs too:

- (133) 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  $\theta$ -roles, but c-select only CPs.

- (134) 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  $\theta$ -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

<sup>13</sup> 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).

## 2. Phrase Structure

---

a CP; that is, there should be no verbs that express these  $\theta$ -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  $\theta$ -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-finite 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 thesis that c-selection can be derived from s-selection fares better when external arguments are concerned. To begin with, the range of categories that serve as external arguments looks somewhat less varied; to a large extent, only NPs and CPs seem to be clausal subjects in English.<sup>14</sup> And second, when a  $\theta$ -role is consistent with either NP and CP, any kind of CP is possible as is an NP:

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

By contrast, when the subject  $\theta$ -role is incompatible with the meanings that CPs yield they are banned from Specifier position:

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

So, let us conclude that only complements are c-selected. This will require weakening (131) to something like (137).

- (137) a. An  $X^0$  c-selects its complements.  
 b. If an  $X^0$  c-selects Y, then it  $\theta$ -marks Y.

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

- (138) a. If a verb has a  $\theta$ -role, then there is exactly one syntactic position to which that  $\theta$ -role is assigned.  
 b. A  $\theta$ -marked position must be occupied by something with the appropriate semantic value.

<sup>14</sup> With the exception of cases like "Under the bed is a slipper," plausibly instances of impersonal constructions with inversion; see Stowell (1981) and Rochemont and Culicover (1990).

- c. A verb *c*-selects its complements.
- d. *c*-selection entails *s*-selection (aka  $\theta$ -role assignment).

The statements in (138a) and (138b) are the THETA CRITERION, whereas those in (138c) and (138d) concern the relation between *c*-selection and *s*-selection which we've just reviewed. The THETA CRITERION insists that for every  $\theta$ -role that some verb has, there will be a unique position occupied by an argument in the sentence holding that verb. (138c) and (138d) 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  $\theta$ -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 the procedure for locating the  $\theta$ -positions for both subject and object arguments is the same, or very nearly so, 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 (139).<sup>15</sup>

(139) THE PROJECTION PRINCIPLE

- i. For  $\alpha$ , a position, if  $\alpha$  is a sister to  $X^0$ , then  $X^0$  *c*-selects  $\alpha$ 's contents.
- ii. If  $\alpha$  *s*-selects  $\beta$ , then  $\alpha$  and  $\beta$  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.<sup>16</sup> 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  $\theta$ -role. We will see,

<sup>15</sup> Chomsky'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*.

<sup>16</sup> X and Y are *sisters* if every phrase including one includes the other.

however, that this problem is only apparent. Once we discover what is truly responsible for assigning the subject its  $\theta$ -role, this problem is resolved (or, depending on how certain particulars play out, mutated into a different problem).

The first part of the Projection Principle is just (138c). It has the interesting consequence of preventing non-arguments from standing in the smallest  $\bar{X}$ . Thus, the Projection Principle not only has the effect 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 phenomena that might 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  $\theta$ -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 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 (140).

- (140) a. If  $\alpha$  modifies  $\bar{N}$ , then  $\alpha$  must be ...  
b. If  $\alpha$  modifies  $\bar{V}$ , then  $\alpha$  must be ...



⋮

The “...” will carry lists of category types.

If this project is successful, then the Phrase Structure 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 (140), 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 phrase structure component of the grammar might have nothing more than (141) in it, where “ $\{\alpha, \beta\}$ ” should be understood as representing both the string  $\alpha + \beta$  and the string  $\beta + \alpha$ .

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

What the morpho-syntactic category of  $\alpha$ ,  $\beta$  and  $\gamma$  are is fully determined by the c-selection properties of  $X^0$  and the language particular principles governing modifier types (i.e., (140)).

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<sup>17</sup> 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

<sup>17</sup> See Greenberg (1963).

## 2. Phrase Structure

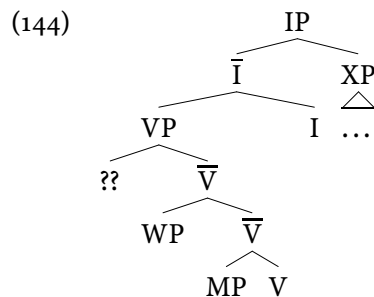
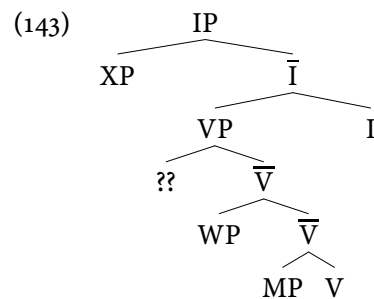
---

(German/Dutch, for example, don't seem to obey this), it does look like this could be true of the underlying arrangements of these constituents.

As a starting point, then, let's take the view that languages linearize their phrasal constituents by way of setting separately the linear order of the immediate constituents of XP and  $\bar{X}$ . This can be achieved by letting languages pick the values "first" and "last" for the terms in (142).

- (142) a. Specifier: [first, last]  
 b. Projection of X: [first, last], modulo (142a)

This connects the linear order of head to complement with the linear order of head to adjunct, which Dryer's work suggests might correct.<sup>18</sup> So, for instance, a language that sets Specifier to "first" and Projection-of-X to "last" will license sentence structures like those in (143), whereas a language that sets both terms to "last" will produce structures like those in (144).

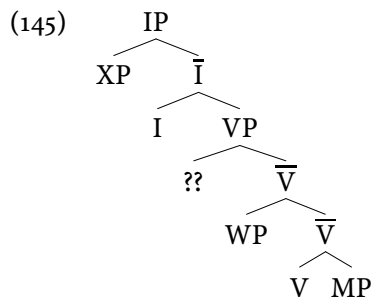


(These phrase markers assume that in the language in question VP is a complement to  $I^0$ , as it is in English.) The categorial values for MP will be determined by the c-selection specification of the verb involved. The categorial values for XP will be determined by the  $\theta$ -role it receives. And the categorial values for WP will be whatever (141) for the language in question allows to modify  $\bar{V}$ s. We haven't yet discovered what sits in the Specifier of VP, so this spot remains ??.

<sup>18</sup> See also Saito (1985) and Saito and Fukui (1998).

The linearization parameters in (142) produce these phrase markers in the following way. Setting Specifier to “first” in (143) linearizes XP and ?? so that they precede  $\bar{I}$  and  $\bar{V}$  respectively. Setting Projection-of-X to “last” makes every other  $\bar{I}$  and  $\bar{V}$ , as well as I and V, follow the phrase they are sisters to. As a consequence WP, MP and VP precede the phrase they are complements to or modifiers of. In (144), by contrast, Specifier is set to “last,” which linearizes XP and ?? so that they follow  $\bar{I}$  and  $\bar{V}$  respectively. As with (143), Projection-of-X is set to “last” in (144) and the consequence for the position of WP, MP and VP is the same.

Restricting the linearization options to just those in (142) blocks certain phrase markers. It blocks languages, for instance, in which the complement to a verb falls on a different side to that verb than does a complement to a noun. That is, it forces languages to unify the linearization of Specifier, Complement and modifier across phrase types. It is not hard to find languages that seem to violate this restriction, but as Greenberg and Dryer find, there is a tendency for languages to avoid this type. Similarly, (142) prevents all languages that put modifiers to one side of the  $\bar{X}$  they modify but put complements to the other side. For instance, phrase markers like (145) are prevented.



This phrase marker linearizes  $\bar{V}$  (a projection of V) “last” relative to WP, but linearizes V (also a projection of V) “first” relative to its complement. Clearly there are languages of this unexpected type; English seems to look precisely like (145).

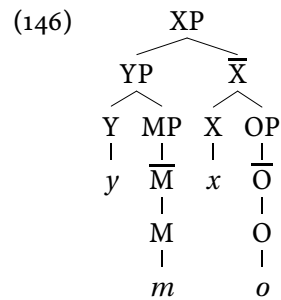
This proposal, then, seems clearly too restrictive. Nonetheless, it will be our starting point. In the chapters that follow we will explore ways of loosening this model so that it is enabled to account for the range of language types we do see without losing the trends in linear organization that Greenberg and Dryer have discovered. What we have now is not yet complete enough to really engage this problem.

In fact, the linearization scheme in (142) is itself not yet complete enough to generate the strings we want to associate with the phrase markers it allows, for example those in (143) and (144). All (142) does is linearize the phrases within a sentence. It does not determine how the strings of words within those phrases are

## 2. Phrase Structure

---

linearized relative to the other phrases. To see this, consider a phrase marker like that in (146) below, in which lower-cased letters should be understood as representing words. This phrase marker arises by imposing the  $\bar{X}$  Skeleton and setting



Specifier to “first” and Projection-of-X also to “first.” What we would like is for this to be sufficient to generate the string  $ymxo$ . Instead, however, all that these settings give us is the information in (147).

- (147)
- a.  $y$  precedes MP
  - b.  $y$  precedes  $\bar{\text{X}}$
  - c.  $x$  precedes OP

What’s required is something to determine that how the information in (147) determines the linear order of  $y$  relative to the words within MP and  $\bar{\text{X}}$  and, similarly, determines the linear order of  $x$  relative to the words within OP. Let’s turn to that now.

Recall that in defining morpho-syntactic category, we entertained the hypothesis that looking at only adjacent terms would be sufficient for defining the relevant substitution classes. In fact the criteria we considered conformed to that constraint. As it happens, in defining phrases we have also obeyed this constraint. As a result, phrases are always strings of adjacent terms and this makes it possible to determine from (147) what the consequent linearization for all the words in (146) is. If the words in MP must be adjacent to each other, then (147a) is enough to know that  $y$  precedes all those words (i.e.  $m$ ,  $x$  and  $o$ ). All that is required is an explicit statement that the words within a phrase are linearized with respect to some other word in the same way that the phrase is. This can be done with (148).

- (148)  $[\alpha, \beta] =_{\text{def.}} \alpha$  precedes  $\beta$ .  
 $\{\delta \alpha, \beta\} =_{\text{def.}} \begin{array}{c} \delta \\ \underbrace{\quad} \\ \alpha \quad \beta \end{array} \text{ or } \begin{array}{c} \delta \\ \underbrace{\quad} \\ \beta \quad \alpha \end{array} .^{19}$

- a. For all words,  $x$  and  $y$ , within a phrase marker, either  $[x, y]$  or  $[y, x]$ .

<sup>19</sup>  $\alpha$  and  $\beta$  are said to be “sisters” in this case, and  $\delta$  is called  $\alpha$  and  $\beta$ ’s “mother.”

- b. 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  $Y$ .

(148b) expresses the hypothesis that phrases are only defined over adjacent items as a constraint on linearization. Because of its apparent universality, (148) is a candidate for a language universal.

On this view, then, the phrase structures of languages are the result of four fixed universals — the  $\bar{X}$  Skeleton, the Theta Criterion, the Projection Principle, and the linearization principles in (148) — plus the following language particular pieces of information:

- (149) a. Specifying the categories of modifiers (i.e., (140)).
- b. Setting the “headedness parameter” (i.e., (142)).
- c. A vocabulary whose lexical items  $c$ -select complements.

There’s a sense, then, in which languages do not actually have Phrase Structure rules. They are merely the epiphenomena that emerge when the various factors of Universal Grammar and language particular information are combined. This theory, if correct, meets the criterion of explanatory adequacy. It provides both inviolable constraints (i.e.,  $\bar{X}$  Theory, the Theta Criterion, the Projection Principle and (148)) and an evaluation metric (i.e., (149) and the language particular vagaries of vocabulary).

Notice how the evaluation metric this proposal embraces is quite different from the “simplicity” metric suggested in Chomsky’s early work. The evaluation metric here involves learning the word-by-word  $c$ -selection requirements and fixing parameter values in the headedness linearization procedure in (149). This proposal has the following form: inviolable constraints come in the form of immutable principles, while the evaluation metric (once lexical idiosyncrasies are removed) consists of principles with a menu of parameters that are set on a language particular basis. Theories that have this general form are said to belong to the “Principles and Parameters” framework. This conception of what explanatory grammars might look like was arrived at by Noam Chomsky and his collaborators in the late 1970’s, and much of the work of the 80’s and early 90’s has this form. In 1981, Chomsky published an ambitious book in which he organized much of the work of that time into a principles and parameters form. This book, *Lectures on Government and Binding*, serves as a rough starting point for much my exposition in these lectures.

In moving from a battery of English specific phrase structure rules to the more explanatory interaction between  $\bar{X}$  Theory and the  $c$ -selection requirements of verbs, language particular settings of modifier types, etc., we have lost some information. Because that transition removed any reference to categories, it is no longer possible to order complements in situations, like (150), where there are more than two.

## 2. Phrase Structure

---

- (150) Sheila put this on the table.  
*compare: \*Sheila put on the table this.*

This information was conveyed in the phrase structure rules by way of referencing category type; we had rules such as (151) for instance which ensure that if a verb is followed by two complements, the first will be the NP.

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

Because the  $\bar{X}$  Skeleton does not have information about category type in it, it is not possible to use the  $\bar{X}$  Skeleton to order complements. Nor would we want to rely on the c-selection requirements of verbs to do this. That would amount to the claim that the order of complements varies as a function of the verb involved. But the fact of English is that no matter what verb is selected, the complements line up in the way that (151) requires.

There must be another component of the grammar which expresses this information. This will be the subject of the next chapter.

# 3

## Positioning Arguments

We've made a whole-scale revision to the information that phrase structure rules provide, placing much of what they formerly said into lexical information and into 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  $\theta$ -role there is exactly one position to which that  $\theta$ -role is assigned.
  - b. For every  $\theta$ -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  $\alpha$ , a position, if  $\alpha$  is a sister to  $X^0$ , then  $X^0$  c-selects  $\alpha$ 's contents.
  - ii. If  $\alpha$  s-selects  $\beta$ , then  $\alpha$  and  $\beta$  are sisters.
- (4) If an  $X^0$  c-selects  $Y$ , then it  $\theta$ -marks  $Y$ .
- (5)
  - a. If  $\alpha$  modifies  $\bar{N}$ , then  $\alpha$  must be AP, PP or CP
  - b. If  $\alpha$  modifies  $\bar{V}$ , then  $\alpha$  must be AdvP, PP, CP or NP

⋮

### 3. Positioning Arguments

---

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 expansions reflect, perhaps, a particular setting of the universal options of Specifier first or last, and  $\bar{X}$  first or last. The linear ordering of  $\{\bar{X}, \text{WP}\}$  and  $\{\bar{X}, \text{QP}\}$  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  $\theta$ -positions are sisters to the terms that assign the  $\theta$ -roles, and that when the term assigning a  $\theta$ -role is a head, that this position is also  $c$ -selected. This leaves the problem of correctly determining the subject's  $\theta$ -position – a problem whose solution we are working towards. The Theta Criterion ensures that for every  $\theta$ -role associated with some predicate, there will be exactly one  $\theta$ -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.

#### 3.1 Expletives and the Extended Projection Principle

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  $\theta$ -roles: if the head of the phrase has  $\theta$ -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  $\theta$ -position for the subject argument, and the Theta Criterion will force an argument into this position. If the  $\theta$ -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  $\theta$ -role associated with a subject argument the presence of something in specifier of IP is obligatory. The verb *seem*, for instance, has only one  $\theta$ -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. 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”  $\theta$ -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  $\theta$ -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  $\theta$ -role. When the verb of some sentence has a

### 3. Positioning Arguments

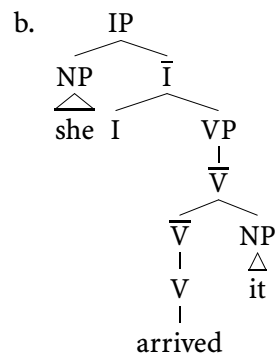
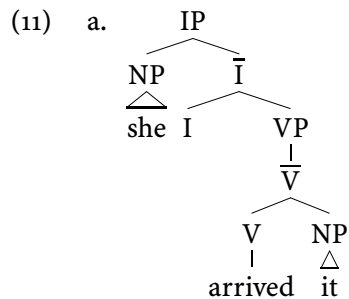
---

subject  $\theta$ -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  $\theta$ -role is associated. As we've seen, they're not found in Specifier positions associated with  $\theta$ -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.<sup>1</sup>

(12)  $\{\bar{X}, \alpha\}$  iff  $\alpha$  modifies  $\bar{X}$ .

<sup>1</sup> Recall that the “{ }” notation expresses hierarchical relations without expressing linear ones.

This bijection limits modifiers to sisters of  $\bar{X}$ s. It is redundant with the Projection Principle, which blocks modifiers from complement position. But it also blocks modifiers from being in Specifier positions, something that would otherwise be available. It also forces sisters to  $\bar{X}$ s to be modifiers, and this is the use we have of it here. 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  $\theta$ -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  $\theta$ -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  $\theta$ -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.

### 3.2 Case Theory and ordering complements

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) and (16).

### 3. Positioning Arguments

---

- (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].

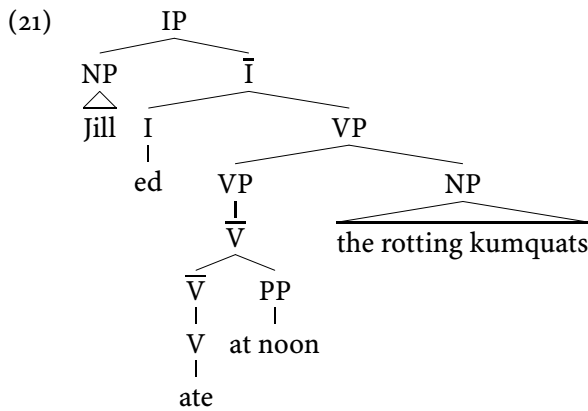
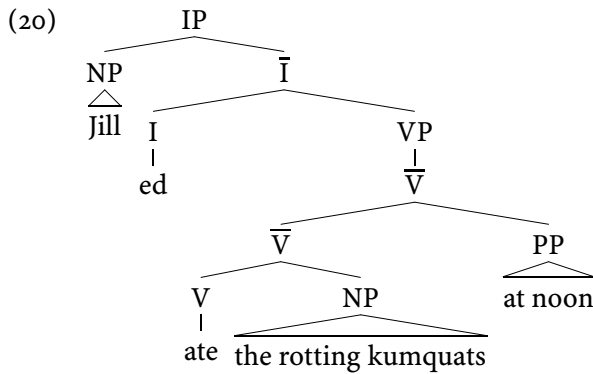
So, we can see already that there is some work to be done by principles that impose a linear order on complements. If we've got everything right up to now, they seem equipped with the power to pervert what appears to be the hierarchical relationship between argument and non-argument.

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 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 processes 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 can be formulated as (19).

- (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 series, then, made up of the two phrase markers shown above. We can see the members of this series 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 series. The other members are produced by the action of transformational rules acting on elements already found in the series. We will eventually have to control the method by which transformations produce additional elements, but in the simple cases we will be concerned with right now, there are only two members of the series: the D-structure and another formed by a transformation acting on that D-structure. We only speak one element of these series. 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 series of parses are called "Derivations."

### 3. Positioning Arguments

---

Consider now how complements are ordered linearly. Recall 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.<sup>2</sup>

(22) NP + 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 Indo-European 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 Indo-European 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 argument NP that sentence contains in a Case marked position.

<sup>2</sup> As a simplifying measure, we consider only finite clauses.

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 — 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 Indo-European 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  $\alpha$  only if  $X^0$  c-commands and is close to  $\alpha$ .
- (26)  $\alpha$  C-COMMANDS  $\beta$  iff:
  - i. every phrase that contains  $\alpha$  contains  $\beta$ , and
  - ii.  $\alpha$  does not contain  $\beta$ .
- (27)  $\alpha$  is KLOSE to  $\beta$  iff there is no more than one phrase that contains  $\beta$  but not  $\alpha$ .

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

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

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)  $\alpha$  assigns Case to  $\beta$  only if  $\alpha$  and  $\beta$  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

### 3. Positioning Arguments

---

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.

Consider, next, 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) Mary will put the book on the table. →  
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 it interacts with another process that moves something to the left edge of IP, as in (33).

- (33) Mary will put the book on the table. →  
On the table, Mary will put the book.

These two processes cannot apply to the same sentence, as (34) indicates.



- (34) a. Mary will put the book on the table →  
b. On the table, Mary will put the book. →  
c. \* 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  $I^0$ .

- (35) 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.

- (36) 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 (37) is a counterexample.

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

But, so far as I know, (37) 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.

### 3. Positioning Arguments

---

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 are. 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 (38).

- (38) 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.

- (39) a. Sean is unhappy that he had to sing.  
b. \* Sean is unhappy that.
- (40) 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 too. 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:

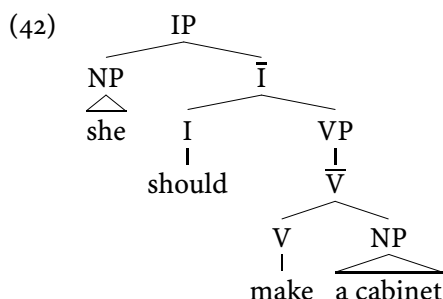
- (41) a. Mary showed Bill the picture.  
b. Bill baked Sally a cake.

How can the second NP in these examples 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.

#### 3.3 Small Clauses and the Derived Subjects Hypothesis

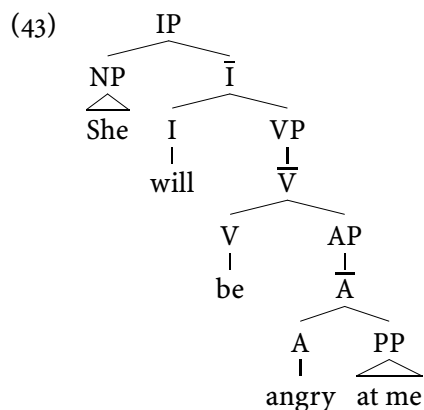
There is a use that we can put the Case filter to that is somewhat related to Stowell's. The Case filter plays a role in solving a problem we have postponed discussing for some time now: How is it that the second clause of the Projection Principle be

correct, given the presence of subjects whose  $\theta$ -role is determined by a verb. The second clause of the Projection Principle, recall, allows  $\theta$ -roles to be assigned under sisterhood only. Subjects are obviously not sisters to the verbs they get their  $\theta$ -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  $\bar{V}$  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 (43) below. 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? One hypothesis is that 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 (44) are examples.

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

### 3. Positioning Arguments

---

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 (45), for example, this position can be occupied by an expletive.

- (45) 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. As (46) indicates, an  $I^0$  is not found in this string.

- (46) 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].

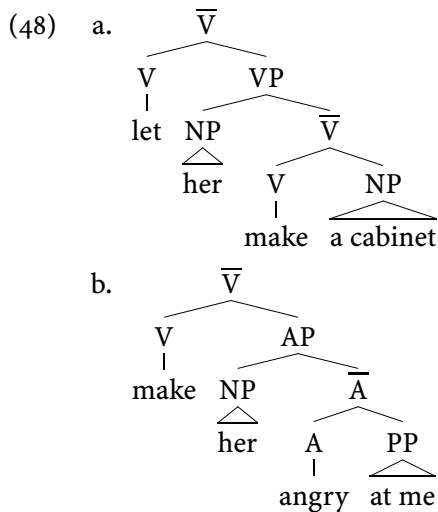
Finally, why don't we think that these NPs are not inside the  $\bar{V}$  or  $\bar{A}$ ? I can't show you a reason for this in the case of  $\bar{A}$ ; but for  $\bar{V}$  we can tell this from the action of the *do so* phenomenon. Recall that *do so* replaces  $\bar{V}$ s; so the contrast in (47) indicates that the NP in (44a) is not in  $\bar{V}$ .

- (47) 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 (44) the parses in (48) on the facing page. 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.<sup>3</sup>

Here, finally, I am able to defend 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,

<sup>3</sup> See Stowell (1978).



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).

Recall that reason we started to look at these examples of small clauses was because they help solve the problem for the Projection Principle that subjects pose. The problem is that subject arguments do not appear to be sisters to the verb, or other predicate, that they get their  $\theta$ -role from. But the Projection Principle requires that  $\theta$ -roles are only assigned under sisterhood. So, let's see now how small clauses help us find a solution to this problem.

Note, first, that the subject of the small clause in (44a) bears the same  $\theta$ -role that the parallel subject in (42) bears. Maybe, then, despite appearances, the position the subject in (42) gets its  $\theta$ -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 (42) were in Specifier of VP to receive its  $\theta$ -role, it still wouldn't be sister to the verb it gets its  $\theta$ -role from.

But wait: who said it gets its  $\theta$ -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  $\theta$ -role from. Indeed, this is just the position that Marantz (1984) advances. He notes that the specific content of the subject's  $\theta$ -role varies as a function not just of the verb, but also of the verb in combination with the material that follows it. This can be seen by considering (49).

### 3. Positioning Arguments

---

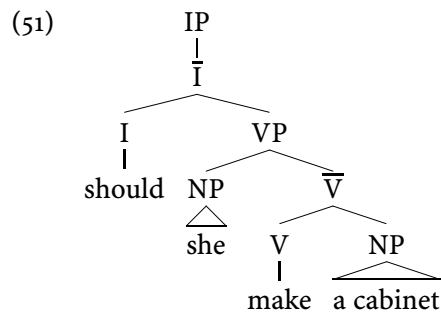
- (49) a. She should make a cabinet.  
 b. She should make an effort.  
 c. She should make some noise.  
 d. She should make nookie.

This doesn't seem to be true of the  $\theta$ -role complements bear, however. It isn't possible to find examples where the specific content of the object's  $\theta$ -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  $\theta$ -role.)

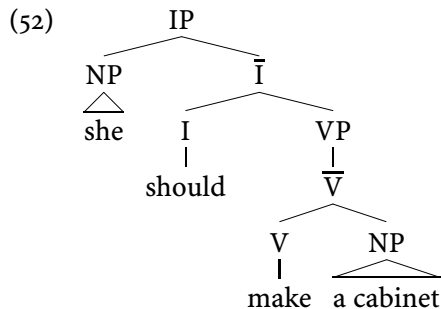
So maybe we should rescue the Projection Principle by allowing certain  $\bar{V}$ s to assign  $\theta$ -roles.

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

This would make the D-structure representation for (42) something like (51).



From this we must find a way of ensuring that the pronounced S-structure in (52) is manufactured.



This solution, then, claims that, in some weird way, subjects are not visibly in the position from which their  $\theta$ -role derives. Instead, they must always undergo a transformational rule that moves them into Specifier of IP. 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 (44b) involves the so-called "Passive" voice, as in (53).

(53) She was made [angry at me].

The  $\theta$ -role that *she* bears in (53) is the same one that it bears in (44b), and yet *she* shows up in a different Case and in a different position in (53). It is a fairly good generalization about the Passive in Indo-European that Passive verbs do not assign Accusative Case, even if the "active" verbs they are related to do. We won't look at Passive in detail, but an informal way of thinking about it is an operation that derives a passive predicate from an active one in the way described in (54).

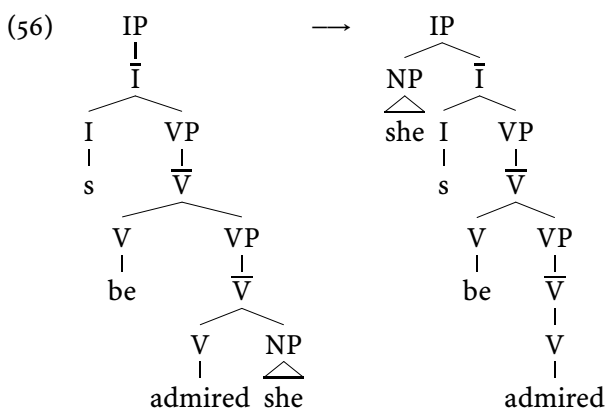
(54) PASSIVE

- a. Add passive morphology to the active verb and embed it under *be*, and
- b. Rob the verb of its Accusative Case, and
- c. Rob the verb of its external  $\theta$ -role.

This takes a garden variety "transitive" verb, like that in (55a), and produces a verb which looks "intransitive," like (55b).

- (55) a. Jerry admires her. →  
 b. She is admired.

On this view of what the Passive does, the D-structure representation of (55b) is the first phrase marker in (56) from which the indicated S-structure is produced.



Because Passive has robbed *admired* of its ability to assign Accusative Case, its object will seek out the Nominative Case marked position to satisfy the Case filter.

### 3. Positioning Arguments

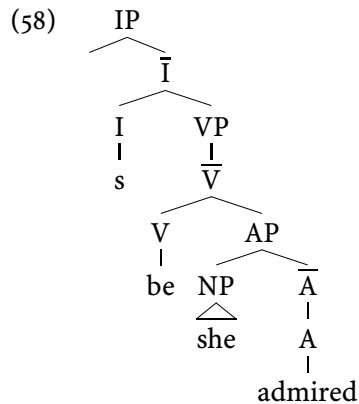
---

This will cause (56) to be mapped onto an S-structure in which *she* is positioned in the Specifier of IP.

There's a confusing aspect to English morphology which clouds what I've just described for (55b). English has a morphological rule by which transitive verbs can be converted to adjectives that look just like passive participles. This process is responsible for created the adjective in (57), for instance.

(57) the admired student

It could be, then, that what we have in (55b) is adjectival phrase headed by *admired*; it might have a D-structure like that in (58).



Over a large class of cases, then, we will not be able to tell whether we are looking at a Passive sentence or a *be*+AP (a so-called “copular construction”) sentence.

In some cases, however, it is possible to tell.<sup>4</sup> This is because the morphological process that derives adjectives from verbs operates on the arguments of the verb. In particular, it makes the “direct object”  $\theta$ -role of a verb the “external”  $\theta$ -role of the resulting adjective. For instance, the adjective in (58) assigns to its Specifier position the same  $\theta$ -role that the verb it was derived from assigns to its complement. This is quite general. We can see this by considering the relation these adjectives have to the nouns they modify. The pre-nominal adjective in (57) is related to the noun it modifies in the same way that the verb it derives from is related to its object. When we figure out how the modification relation works semantically, we will need to derive the fact that the modification relation for adjectives roughly matches the relation these adjectives have to their subjects in the copular construction. We can use this correlation, then, to see which  $\theta$ -role is “externalized” by the adjective formation process. What we find when we look is that, in fact, it’s always the “direct

<sup>4</sup> This point is made in Wasow (1977), where the way of resolving the problem we will adopt is also proposed.



object”  $\theta$ -role that gets externalized. In each of the cases in (59), for instance, the modified noun is related to its adjective in the same way that the direct object is to the verb.

- (59) a. Sally introduced the man to Sean.  
           the introduced man  
       b. Sally placed the book on the table.  
           the carefully placed book  
       c. Sally baked the cake for Sean.  
           the baked cake

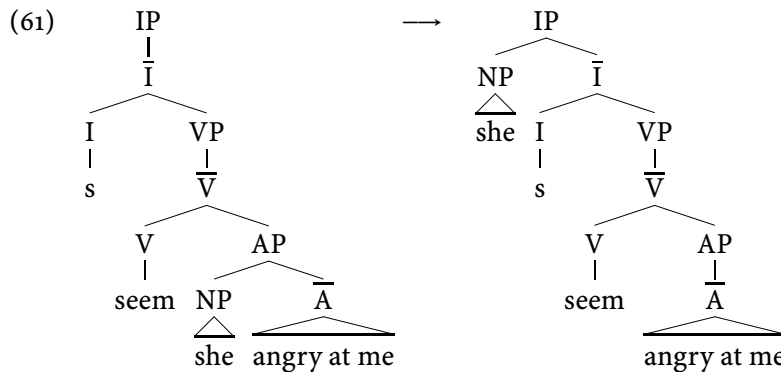
In no case is this relation like the one that holds between the verb and its subject argument or its “indirect object” argument.<sup>5</sup>

Because of this restriction on the adjective formation process, we can be certain that the example in (53) is a Passive and not an adjective. This is because the argument that shows up as the nominative argument in this example is not an argument of the verb/adjective at all. It is the subject of the complement of the verb. If *made* were an adjective in this example, then it’s the complement itself that would have appeared as subject.

There are verbs which behave much like the Passive verb in (53), but without needing to be Passivized. An example is (60).

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

Again, *she* is the subject of *angry at me*, not *seems*. And yet, just as in (53), the subject of *angry at me* shows up in a higher Specifier of IP; this example seems to require the derivation in (61).



Another example like (60), which we shall have occasion to encounter frequently, is (62).

<sup>5</sup> For more details about this process, see Levin and Rappaport (1986)

### 3. Positioning Arguments

---

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

In this example too, *she* receives its  $\theta$ -role from the lower clause: *she* is the external  $\theta$ -role bearer of *angry at me*. In this example, however, *angry at me* is embedded within an IP headed by *to*.

In all of these cases, then, we seem to need to let arguments appear in positions distant from their  $\theta$ -role assigners. Whatever we come up with for these examples can be applied to the simpler case in (51). That is, holding the view that external  $\theta$ -role bearers start out in Specifier of VP and moves to Specifier of IP, in a garden variety simple sentence, is not adding anything to the grammar that isn't needed already. It is a simple version of the cases we've just reviewed.

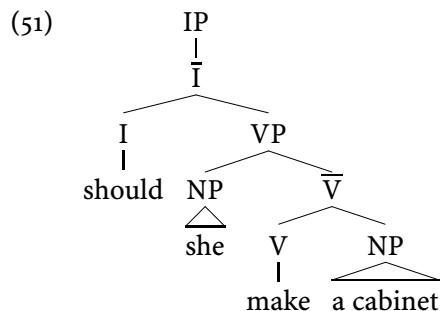
How are we to account for these situations? It is common to follow Ross's technique for solving cases where the linear relationships we'd expect from the Projection Principle fail to materialize. We speculate that the subject in these situations undergoes a Transformation which moves the argument into the Specifier of IP we see it in. Let's call this Transformation ARGUMENT MOVEMENT, and give it the preliminary formulation in (63).

(63) ARGUMENT MOVEMENT

Move an XP from a  $\theta$ -marked position into a Specifier position.

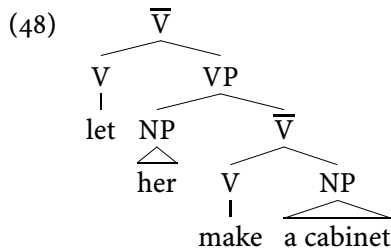
If this is the rule that produces the derivations we've been examining, then what we need to do now is understand how to force it to apply in these situations (and prevent it from applying too broadly).

Let's return, then, to (51) (repeated below) which we are considering to be the D-structure for *She should make a cabinet*.



The Theta Criterion and Projection Principle (among other constraints) are responsible for placing *she* in the Specifier of VP. What is it that prevents *she* from remaining in this position? Why must ARGUMENT MOVEMENT apply to form an S-structure in which *she* is spoken in Specifier of IP?

It is instructive to compare (51) to the D-structure representation in (48) for the matching VP in *I let her make a cabinet*, repeated below.



Again, the Theta Criterion and Projection Principle are responsible for positioning *her* in Specifier of VP in this parse. But in this scenario, Specifier of VP is also the S-structure position for *her*. ARGUMENT MOVEMENT is not required to move *her* into a different surface position in this case.

What distinguishes these cases? One thing is the Case morphology borne by the pronoun. In (51) it is nominative (*she*) and in (48) it is accusative (*her*). Indeed, the system of Case assignment we have formulated distinguishes (51) and (48) in the way desired. We would expect the subject to be unable to remain in Specifier of VP in (51) because this is not a Case marked position. Accusative Case is not assigned by *make* to this position because *make* does not c-command it. Instead, the only way *she* can satisfy the Case filter is by being moved into a position that assigns Case, and the only such position in (51) is the nominative Specifier of IP position. By contrast, *make* does c-command *her* in (48), and this is what allows *her* to remain in this position and satisfy the Case filter. One hypothesis, then, is that ARGUMENT MOVEMENT is triggered to satisfy the Case filter.

If that is correct, it is possible to diagnose certain examples as Case filter violations by seeing how they are related to grammatical examples that employ ARGUMENT MOVEMENT. Consider, for instance, (64).

- (64) a. \* It seems [her angry at me].  
 b. \* 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. In fact, however, the only grammatical sentences with the meaning aimed at in (64) is (65).

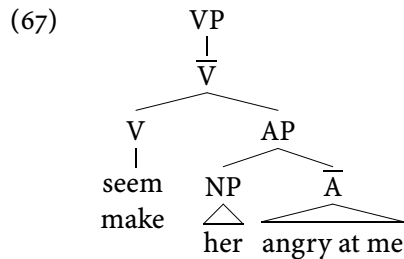
- (65) a. She seems [angry at me].  
 b. A child seems [angry at me].

This would follow if *her* and *a child* do not stand in a Case marked position in (64). It's interesting to compare (64) to (66), both of which have the very similar VPs shown in (67).

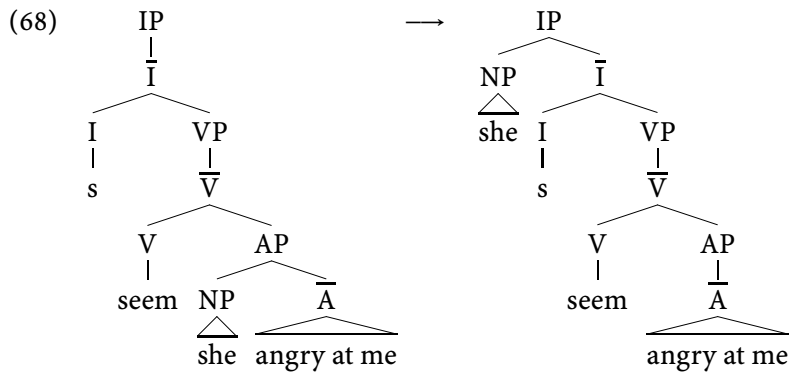
- (66) It makes [her angry at me].

### 3. Positioning Arguments

---



This comparison shows us that the failure of the pronoun to get Case in (64) cannot be because it isn't governed by *seem*. Instead, the difference must devolve to something that distinguishes *seem* from *make*. One hypothesis is that *seem* is simply not equipped with the accusative Case assigning ability that *make* is. Thus, the ungrammaticality of (64) traces back to the fact that the only term that governs the pronoun does not assign Case, and a Case filter violation results. The grammatical outcomes in (65), then, involve the derivation in (68).

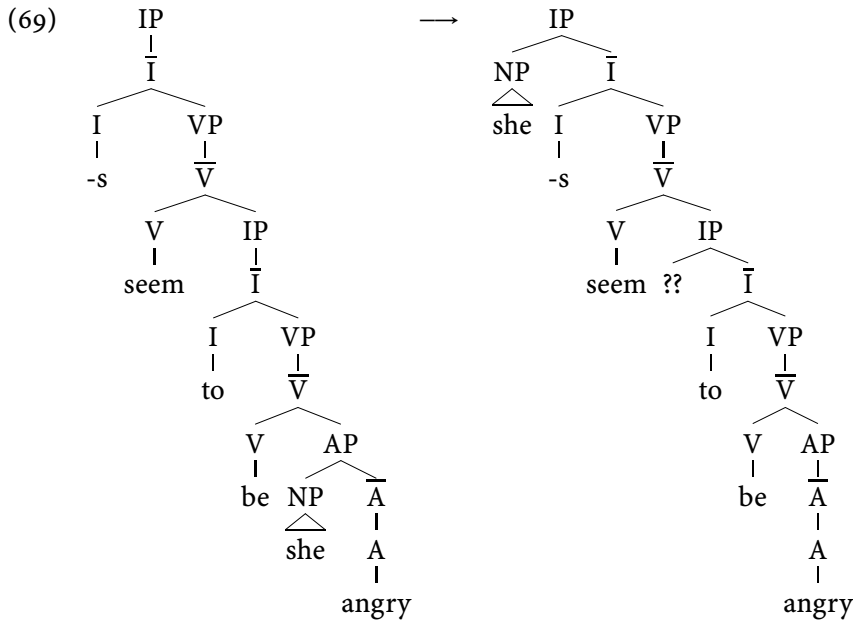


We should give a similar derivation to (62), repeated here.

(62) She seems to be angry.

This derivation is shown in (69) on the next page. Notice that in this example, however, 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 contains a parse in which it is satisfied. Let's adopt his solution for the moment; we'll return momentarily to some evidence in support of it.



### 3.4 PRO and Control Infinitives

The example in (62) has a superficial similarity to (70), but there is an important difference between them.

(70) She tried [to be angry]

In (62), the count of  $\theta$ -roles and arguments matched, as expected. But in (70), this isn't the case. Both *tried* and *angry* assign an external  $\theta$ -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 (71a) doesn't have the meaning that (71b) does.

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

So, either we should find another explanation for the fact that (71) illustrates, or we should find a way of resolving (70) with the Theta Criterion.

Note also that (70) seems to be a violation of the Extension to the Projection Principle. Moreover, it's not just that (70) 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  $\theta$ -role that *angry* assigns, the result is ungrammatical.

### 3. Positioning Arguments

---

(72) \*She tried [him to be angry].

Infinitives with this peculiar mix of properties are called control clauses, another example of which is (73).

(73) Sally prefers to eat chocolate.

Again, as with (70), there is expected to be a subject argument for *eat chocolate* here. Interestingly, however, in this case such an argument can be found if the conditions are changed slightly:

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

We think that the *for* in (74a) 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.

- (75) 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 (75b) derives from the fact that there is only room for one phrase between the *it's* and the *that* of these clefts, and in (75b) two things, a PP and an infinitive, have been shoved into this spot. Thus the grammaticality of (75a) would seem to argue that we do not want to parse the boldfaced 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 (74) 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:

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

Furthermore, the constraints on this optionality, which we can credit to, say, a C<sup>0</sup> Deletion rule, 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 (77). A parallel constraint on *for* ellipsis is indicated by (78).

- (77) a. It's that he eats chocolate that Sally said.  
b. \* It's he eats chocolate that Sally said.  
(78) 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 (74) able to be expressed, whereas the one in (70) 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*.

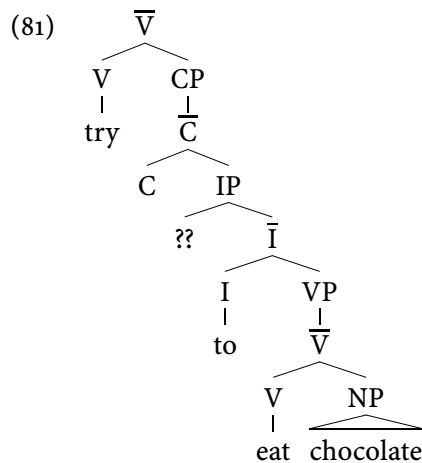
(79) \*She tried [for him to eat chocolate].

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.)

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

(80) Sally tried the task.

So the reason that (72) 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 (81) on 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 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 (82)

### 3. Positioning Arguments

---

shows, there is a perfect correlation with Case marking concerning whether or not the Theta Criterion is satisfied.

(82) \*Sally prefers [for to eat chocolate].

This correlation can be expressed this way with (83).

- (83) a. If the Specifier of IP is Case marked, then the external  $\theta$ -role bearer must be expressed in it.  
b. If the Specifier of IP is not Case marked, then the external  $\theta$ -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  $\theta$ -role to assign, the presence of the missing subject creates ungrammaticality!

- (84) 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 (83), and the facts above, is to imagine that there is an element that bears the external  $\theta$ -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.

(85) 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  $\theta$ -role. We will have more to say about where PRO is found in a few paragraphs.

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  $\theta$ -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  $\theta$ -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 left the last section with — is to adopt the kind of solution that Ross did for the similar problem with complement positions. That is to rely on a Transformation, repeated here.



(86) ARGUMENT MOVEMENT

Move a phrase from a  $\theta$ -marked position to an empty Specifier position.

### 3.5 Evidence for Argument Movement from Quantifier Float

Sportiche (1988) presupposes this solution to the argument spread phenomenon, and argues that Argument Movement also plays a role in relating an external  $\theta$ -role bearer in Specifier of IP with a  $\theta$ -marked position in Specifier of VP. Thus, his article is a defense of the solution to the problem of external  $\theta$ -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 (87).<sup>6</sup>

- (87) 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:

- (88) 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 judgments for some of the French and English examples.

I will try to pull 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.

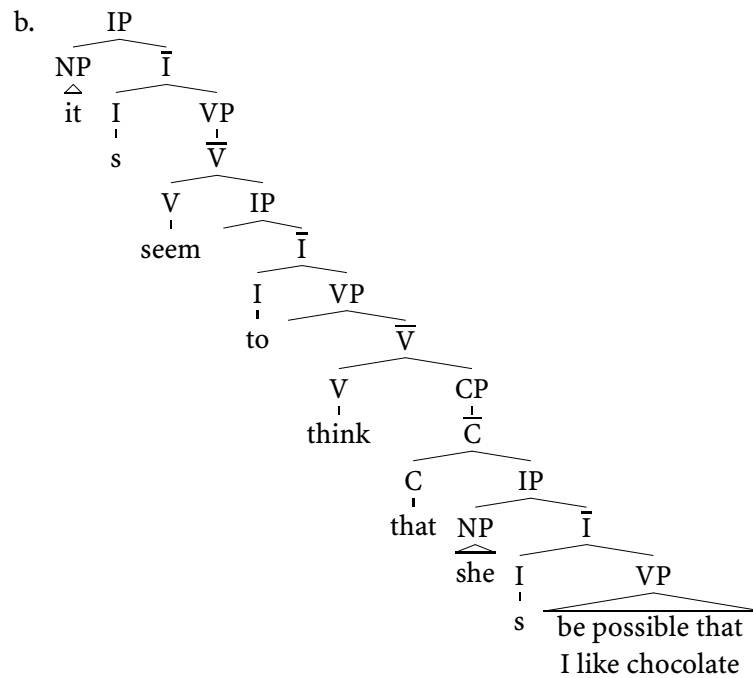
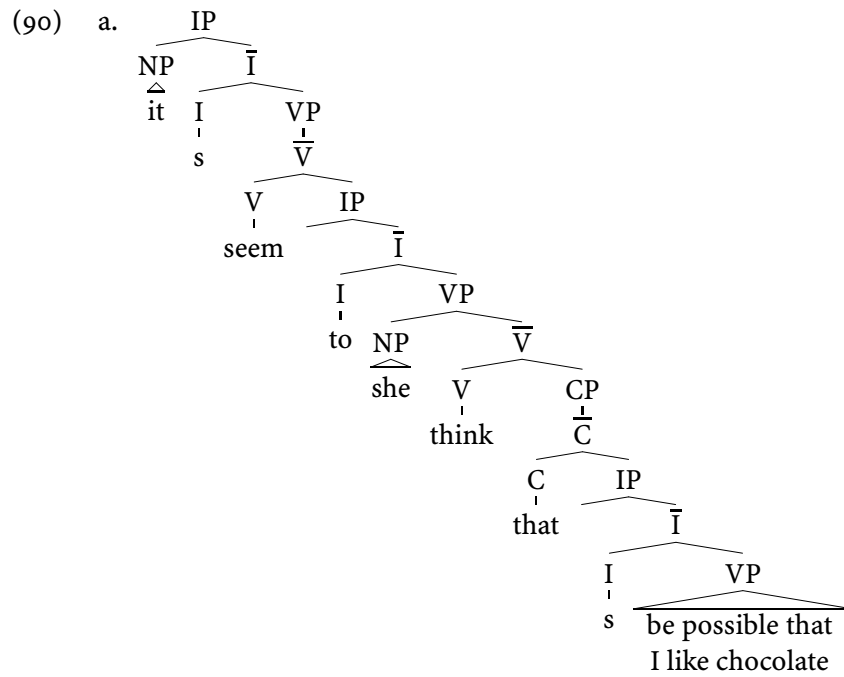
Sportiche’s argument relies on the observation that the relation between the Floated Quantifier and the argument it is related to mimics the relation that is found in cases where an argument is spread across two positions. Thus, for example, this discontinuous relation is constrained in such a way that the position the argument is found in always c-commands the other,  $\theta$ -marked, position. Examples like (89) are ungrammatical.

- (89) 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].)

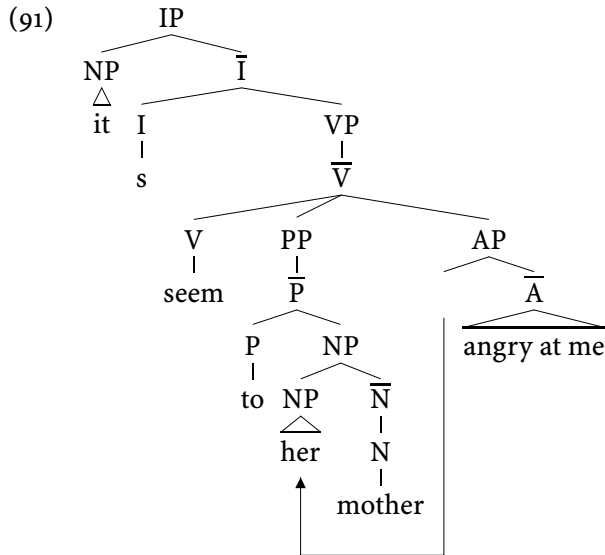
<sup>6</sup> Miyagawa (1989) provides a similar analysis for Quantifier Float based on Japanese data.

### 3. Positioning Arguments

At present, our grammar would allow (89a) to arise through the derivation in (90).



It would also allow (89b) to be created with the derivation in (91).



In both cases, what distinguishes the bad derivations from the grammatical ones is that the argument NP has moved to a non-c-commanding position. If we express the argument spreading phenomenon in terms of Argument Movement, these facts would require that we place the following constraint on it.

- (92) When an argument moves from  $\alpha$  to  $\beta$ ,  $\beta$  must c-command  $\alpha$ .

A similar constraint is seen in Q-Float. It isn't possible for (93b) to have an interpretation like that which (93a) gets, presumably because it is not possible for Q-float to relate these two structures. And because (93b) can have no other interpretation – there are no other plural NPs for *tous* to be construed with – the sentence is ungrammatical.

- (93) 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 ungrammaticality of (93b), then, can be credited to the fact *ces livres* fails to c-command *tous*.

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 (94).

### 3. Positioning Arguments

---

- (94) \*Marty was preferred [ to be angry at me].  
( $\approx$  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, remembering that control infinitives are CPs, we can constrain it with something like (95).

- (95) An argument cannot move from  $\alpha$  to  $\beta$  if  $\alpha$  is within a CP and  $\beta$  is not within that CP.

The Q-Float relation observes a similar constraint:

- (96) 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 semantic 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 (97).

- (97) 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 (98) indicate that argument spread is implicated in the positioning of subjects of simple monoclausal examples.

- (98) The kids should all eat.

More particularly, *all* in this example marks the position that the  $\theta$ -role which *the kids* bears is assigned. Because *all* in (98) 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  $\theta$ -roles be assigned by  $\bar{V}$ s to their sister positions. Thus it supports our solution to the conditions under which external  $\theta$ -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:

- (99) 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*.

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

### 3. Positioning Arguments

---

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 (101), in which it relates NPs, rather than merely taking one as its restrictor.

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

So let's set *each* aside. See Safir and Stowell (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 (102).

(102) 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 (103) indicates.

(103) 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:

(104) 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 (105).

(105)  $\bar{X}$ s cannot be moved.

The match in (99), then, can be explained by way of (105), 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.

(106) 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 (107) and (108).

- (107) 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)
- (108) 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 (109) and those in (110).

### 3. Positioning Arguments

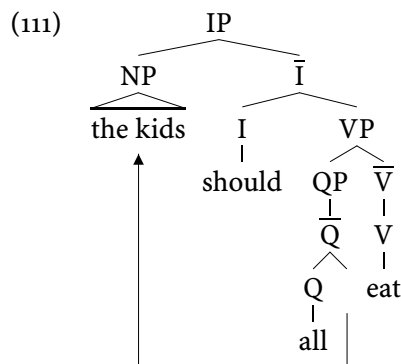
---

- (109) 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)
- (110) 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.

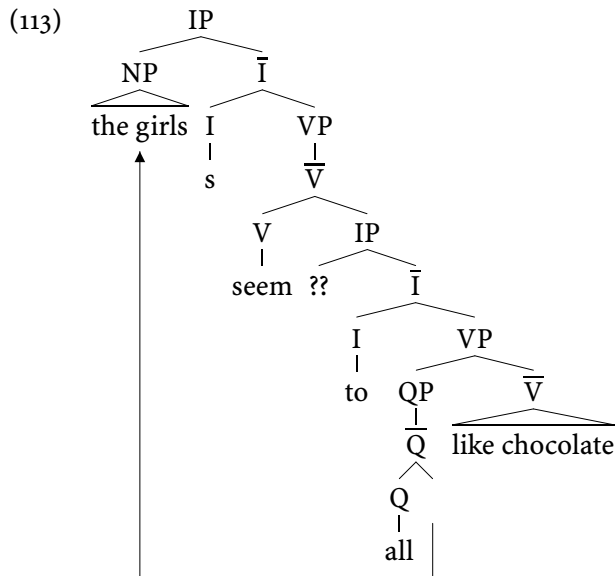
If we call the phrases that put *all* and *both* in construction with an NP, QPs, then a derivation for a simple case like (98) would then look like (111), under Sportiche’s analysis.



A slightly more complex case, one involving two clauses, as for instance occurs with (112), gets the derivation in (113).

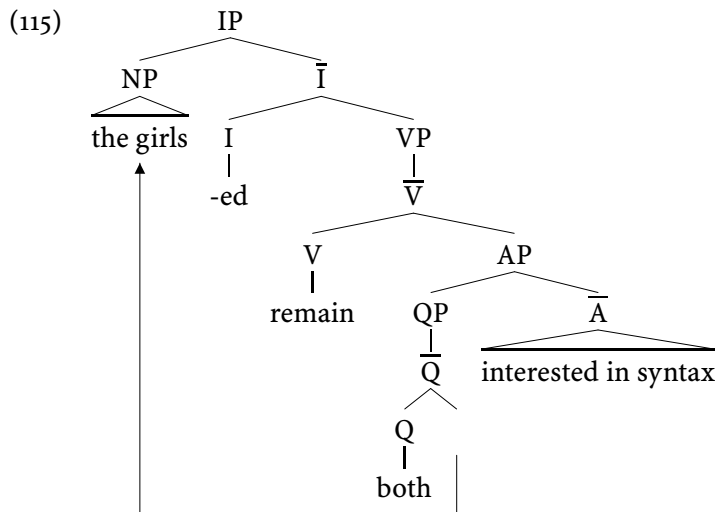
- (112) The girls seem to all like chocolate.





And an example involving a “small clause,” such as (114) gets a derivation like that in (115).

(114) The girls remained both interested in syntax.



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

### 3. Positioning Arguments

---

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  $\theta$ -marked position to a Case marked position. So in (116), for example, it appears that *the children* has made an intermediary stop in the Specifier of the VP headed by *have*.

(116) The children might all have eaten chocolate.

And in (117) we see that *the girls* can make a stop in Specifier of the infinitival IP.

(117) 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.

#### 3.6 Towards a typology of infinitive types

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 (94), is derived. The condition on Argument Movement in (95) forbids movement out of CPs, recall, and this is what we credited the ungrammaticality of (94) 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 (118a) and (118b).

- (118) a. \* I tried [<sub>CP</sub> [<sub>IP</sub> him to be happy]].  
b. I believe [<sub>IP</sub> 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 (118b) is fine, but passivizing (118a) is ungrammatical.

- (119) a. \* He was tried [<sub>CP</sub> [<sub>IP</sub> he to be happy]].  
 b. He was believed [<sub>IP</sub> he 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:

- (120) 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:

- (121) \*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  $\theta$ -role bearer we have discovered in control infinitives. In describing where we have discovered this item, PRO, we came up with (85).

- (??) 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  $\theta$ -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 (122).

- (122) a. She prefers [<sub>CP</sub> [<sub>IP</sub> PRO to drink kava]].  
 b. \* She prefers [<sub>CP</sub> for [<sub>IP</sub> PRO to drink kava]].

### 3. Positioning Arguments

---

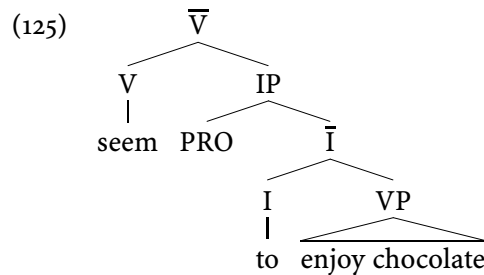
If *for* is a Case assigner, then the position PRO sits in in (122b) has to be a Case marked position, whereas the one it sits in in (122a) might manage not to be. (For instance, we might imagine that *for* deletes in the derivation leading to (122a) before it assigns its Case.)

That there is something wrong with the description in (85) 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 (123) doesn't have an interpretation like (124).

(123) It seems [ to enjoy chocolate].

(124) It seem that someone enjoys chocolate.

The only way for (123) to be grammatical is for *it* to be understood as the argument bearing the external  $\theta$ -role assigned by *enjoy chocolate*; (123) is ungrammatical if *it* is understood to be an expletive and the bearer of the external  $\theta$ -role is PRO. Something, then, must prevent PRO from finding itself in the position indicated in (125).



(85) won't do this, since the Specifier of the IP in (125) is, as required, not Case marked. In fact, (125) 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:

(126) 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 (127) that is not true in the two grammatical examples in (128), is that PRO is close to a word.

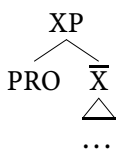
- (127) a. \* It seems [<sub>IP</sub> PRO to enjoy chocolate].  
 b. \* Sal prefers [<sub>CP</sub> for [<sub>IP</sub> PRO to drink kava]].
- (128) a. Jerzy tried [<sub>CP</sub> [<sub>IP</sub> PRO to enjoy chocolate]].  
 b. Sal prefers [<sub>CP</sub> [<sub>IP</sub> PRO to drink kava]].

Because of the presence of the CP (and  $\bar{C}$ ) in (128), there are two many phrases “between” PRO and the higher verb. In (127a), by contrast, there is only one phrase “between” PRO and *seem*; and similarly, there is only one phrase “between” PRO and *for* in (127b). We could describe this contrast, then, with (129).

- (129) PRO may not be c-commanded and klose 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 (129) 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 (129).

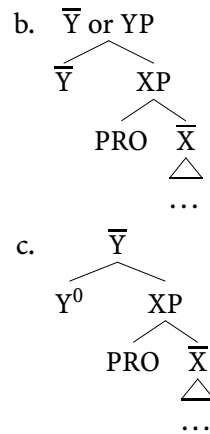
For the time being, then, let’s adopt (129) as part of our description of where PRO can be. In fact, (129), when combined with some of the other principles we’ve seen, will dramatically restrict the distribution of PRO, and comes very close to doing everything that (85) did. If PRO does not have a meaning 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 klose to the word that selects it.<sup>7</sup> Thus, (129) will force PRO to occur only in Specifier positions. Further, (129) will allow PRO to exist in a Specifier position in just the configurations in (130), where  $Y^0$  is a position that is not occupied by a lexical item.

- (130) a. 

<sup>7</sup> (129) 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, (129) would let PRO exist as the objects of deleted verbs.

### 3. Positioning Arguments

---



When the phrase whose Specifier sits in is a complement, as in (130c), 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, (129) lets PRO exist only in the Specifiers of unembedded phrases (i.e., (130a)), and the Specifiers of phrases in adjunct or Specifier position (i.e., (130b)), or in the Specifiers of embedded IPs. We've not seen PRO in the positions that (130b) describes, but we will. It turns out that PRO is not possible in the Specifier position of unembedded phrases, as in (130a). Because the only unembedded clauses we are likely to see are finite IPs, we'll need (131) in addition to (130), then.

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

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 (132).

(132) \*Jerzy said [<sub>CP</sub> [<sub>IP</sub> PRO likes chocolate]].

We'll have a chance to reexamine this way of describing where PRO exists; but let's adopt (130) and (131) 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 (133).

- (133) a. \* It appears [to *Mary* 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 [<sub>CP</sub> that [<sub>IP</sub> it seems [<sub>IP</sub> to *Mary* like chocolate]]].  
           (*compare*: It appears that *Mary* seems to like chocolate.)

In (133a), 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 (which is shown in shaded font). And the S-Structure in (133b) is prevented by the constraint that blocks relocating an argument out of a CP.

### 3.7 Constraints on Argument Movement and the typology of verbs

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 (134).

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

This constraint has been formulated in a variety of ways; (135) will do for the moment.

- (135) Move  $\alpha$  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 (136) on the next page.

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  $\theta$ -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  $\theta$ -marked, but non-Case Marked position, just when an expletive can stand in the Case Marked

### 3. Positioning Arguments

---

- (136) ARGUMENT MOVEMENT  
Move a term from  $\alpha$  to  $\beta$   $\beta$  an empty position licensed by  $\bar{X}$  Theory,<sup>8</sup>  
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  $\beta$ .

position we would have understood A-Movement to involve. There are very narrow conditions on this alternative, as we'll see soon, but some examples which fall within those narrow conditions are (137).

- (137) 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 (138).

- (138) 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 (139).



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

This reformulation will require us to re-craft the Case filter as well, so that it can be satisfied by an A-Chain. Let's adopt (140).

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

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, 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 (141) indicate, and the sorts of predicates surrounding the argument seem to matter too, as the ungrammaticality of (142) shows.

- (141) a. \* There might be Frank walking up my neck.  
       b. \* There seems to be Frank walking up my neck.
- (142) 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.)  $\theta$ -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

### 3. Positioning Arguments

---

can create. They won't allow verbs to exist which create conflicts which Argument Movement can't fix.

Verbs set up these situations by creating  $\theta$ -marked and Case marked positions. Thus, for example, we have seen that there are verbs which assign one external  $\theta$ -role, and verbs which assign an external and an internal  $\theta$ -role and Accusative Case. These are sometimes called intransitive and transitive verbs, respectively; examples are in (143).

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

A special kind of transitive verb are exemplified by *believe*, *consider* and *make*, which assign their Accusative Case to something different than the argument they assign an internal  $\theta$ -role to. As we've seen, these verbs can take a "clausal" complement — sometimes these are "small clauses" — and assign their Accusative Case to an NP within this clause.

- (144) a. She believes [<sub>IP</sub> him to be unhappy].  
b. She considers [<sub>AP</sub> him happy].  
c. She made [<sub>VP</sub> him dance].

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

- (145) Sally appears [ to like kiwis].  
Sally seems [ happy].

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

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

There is evidence that some verbs which might otherwise look like intransitive verbs fall into this last class too. For instance, when *appear* *c*-selects not a clause, as it does in (145), but an NP, it arises in one of the two forms shown in (147).

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

Burzio (1986), who produces the first systematic arguments on behalf of these two classes of single argument verbs, uses the terms *unergative* and *ergative* to distinguish them. Others, notably David Perlmutter who is the co-discoverer of this distinction, have used the term *unaccusative* for what Burzio calls ergative. verbs. Let's use the term *intransitive* as a label for either single argument verb, with these two sub-classifications.

(148) Intransitives

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

There are a wide assortment of syntactic phenomena that are sensitive to the distinction between these two sorts of intransitive verbs. We will encounter a few of them in the lectures that follow. In English, one of the phenomena that confirms the picture that there are intransitives that have a single internal argument as well as those that have a single external argument comes from the adjective formation process we briefly discussed in the previous chapter. Recall that this rule creates from a verb an adjective whose external  $\theta$ -role is the same as that assigned by the verb to its “direct” object. This process, then, should only be able to apply to verbs that have a direct object  $\theta$ -role, and indeed it is blocked for a large range of intransitive verbs as a consequence.

- (149)
- a. \* the danced man
  - b. \* the ran dog
  - c. \* the slept woman
  - d. the cried child

But there are a small class of intransitives which are able to be adjectives by this process, and these are the ergative or unaccusatives.

- (150)
- a. the fallen leaves
  - b. the recently arrived package

Let’s consider, then, the space of verb types that we might expect to find and compare that to what we have found so far.

	$\theta$ -roles	Accusative Case	No Accusative Case
(151)	External, no internal	??	unergatives ( <i>sleep</i> )
	No external, internal	??	ergatives ( <i>appear</i> )
	external and internal	transitives ( <i>like</i> )	??
	no external, no internal	??	??

If assigning an external  $\theta$ -role, an internal  $\theta$ -role and assigning Accusative Case are independent properties that verbs have, then this table expresses all the possible ways in which we might expect these properties to combine. As can be seen, there are quite a number of verb types that we haven’t yet seen, but which we should expect to exist.

Are there verbs yet to be discovered that fill these various cells? In some cases, the properties combined make for verbs that are rather difficult to discover. Consider, for instance, a verb that assigns an external  $\theta$ -role and Accusative Case, but

### 3. Positioning Arguments

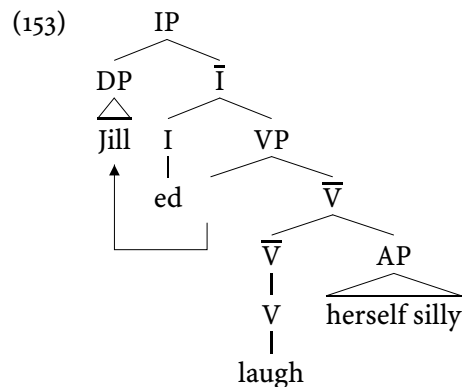
---

assigns no internal  $\theta$ -role (a verb that would fit in the cell in the top row, first column). It will be very difficult to discover verbs of this sort, even if they should exist, because without a complement there will be nothing to bear the Accusative Case that such a verb would assign. The only way to see such a verb would be in cases where we might find a non-complement to which, or into which, Accusative Case could be assigned. One candidate, perhaps, for this situation are cases such as (152).<sup>9</sup>

(152) Jill laughed herself silly.

It is likely that the small clause, *herself silly*, is not a complement to *laugh*; it does not, in any case, refer to something that is involved in the event, or action, that *laugh* denotes. If the Accusative Case on *herself* comes from *laugh* — and where else could it come from — then *laugh* is a verb of the sort we are in search of.

It should be noted that this analysis of (152) is at odds with some of the rest of the grammar we have developed. If *herself silly* is not  $\theta$ -marked by *laugh*, then the Projection Principle is going to require it to be a sister to the  $\bar{V}$  that *laugh* projects, as indicated in (153).



But in this configuration, *laugh* will not c-command *herself*, and this is a requirement on Accusative Case assignment. Indeed, if the c-command requirement on Case assignment is correct and the Projection Principle's placement of complements is too, then these will conspire to prevent verbs of the sort we are searching from ever being found. If *laugh* genuinely is such a verb, then these parts of our grammar will need adjustment. This is work for the future.

<sup>9</sup> My thanks to Angelika Kratzer for suggesting that I use the adjective formation rule as a diagnostic for unaccusatives, and for offering this construction as an example of this class of verb. See Carrier and Randall (1992) for some discussion of this latter construction.

Consider now verbs that assign neither an external nor an internal  $\theta$ -role: the class of verbs that would fill the cells of the bottom row in (151). Do these verbs exist? A candidate are verbs such as *rain*:

(154) It rains.

If the *it* in this example is not an argument, then here is a verb that assigns no  $\theta$ -role. It's conceivable, on the other hand, that *it* is an argument in this example. There might be something – Thor perhaps? – that is doing the raining. Whether or not verbs of this sort exist I'll leave open: let's leave question marks in these cells then.

What of the other two categories of missing verb? Are there verbs which support no external  $\theta$ -role, but do assign an internal  $\theta$ -role and Accusative Case? And are there verbs that assign both external and internal  $\theta$ -roles, but no Accusative Case? To date, there are no verbs with these properties that have been discovered. At present, then, we can update the table in (151) to (155).

	$\theta$ -roles	Accusative Case	No Accusative Case
(155)	External, no internal	<i>laugh?</i>	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	??	??

Burzio discovered the two gaps in this paradigm where there appear to be no verbs, and formulated generalizations which describe these absences.

(156) BURZIO'S GENERALIZATION

- a. If a verb assigns Accusative Case, then it determines an external  $\theta$ -role.
- b. If a verb determines an external  $\theta$ -role (and an internal  $\theta$ -role?), then it assigns Accusative Case.

Why haven't we found verbs like these? Burzio's generalizations might reflect a relationship between Accusative Case and  $\theta$ -role assignment for which we should find a source.

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 (157).

(157) [<sub>IP</sub> should [<sub>VP</sub> Smith V Jones ]].

If V assigns these two  $\theta$ -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

### 3. Positioning Arguments

---

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:

- (158) [IP – should V<sub>1</sub> [XP – [VP Smith V<sub>2</sub> Jones]],  
where both “–” positions are Case marked.

Now the first part of Burzio's Generalization tells us that V<sub>1</sub> cannot assign Accusative Case. If it did, then it would also assign an external  $\theta$ -role, and that's going to bring the count of things that need Case to 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 V<sub>1</sub>. So far, the only ways we have seen for this to be done are if XP is in a CP:

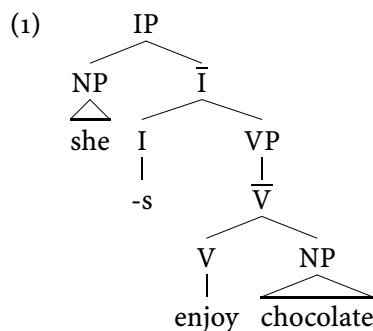
- (159) [IP – should V<sub>1</sub> [CP for [IP – to [VP Smith V<sub>2</sub> Jones]]]].  
[IP – should V<sub>1</sub> [CP that [IP – I<sup>0</sup> [VP Smith V<sub>2</sub> Jones]]]]

But now, there is no way for the arguments of V<sub>2</sub> 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. Thus, 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.

# 4

## Verb Movement

One problem 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 it is this process that is examined in our reading: Pollock (1989).

### 4.1 The “Classic” Verb Movement account

Let’s call the type of argument that Chomsky and Emonds use a “Correlation Argument,” because it’s based on the correlation that holds between a verb’s syntactic position and its inflection. The generalization, in other words, that accounts like these hope to derive is that there appears to be a dependence between a verb’s syntactic position and its inflectional class. This can be seen from the relative positions of auxiliary verbs in English finite clauses. An auxiliary verb cannot follow *not*, *so*, *too* (sometimes called “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.

But it can when if it is not inflected with tense/subject-agreement morphology, as in (3).

- (3) a. Gary should not have eaten.
- b. Gary should not be eating.
- c. Gary will too/so have eaten.
- d. Gary will too/so be eating.

In these contexts we find that the polarity item cannot be found preceding the modal.

- (4) a. \* Gary not should have eaten.
- b. \* Gary not should be eating.
- c. \* Gary too/so will have eaten.
- d. \* Gary too/so will be eating.

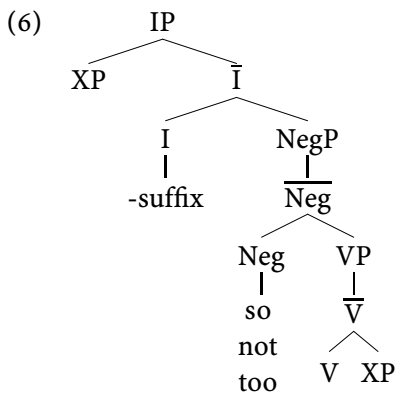
These data lead us to conclude that polarity items lie between  $I^0$  and VP. And from this we can conclude from the contrast between (2) and (3) that verbs are in  $I^0$  when they bear the inflection that resides there. There is, in other words, a correlation between a verb’s inflectional class and its syntactic position relative to polarity items.



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, as (5) shows.

- (5) 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.

A Verb Movement 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 verbs. Thus, simple finite clauses are assigned an underlying representation like that in (6), and clauses with a complex tense (i.e., an auxiliary verb), could have the underlying representations in (7) on the following page.

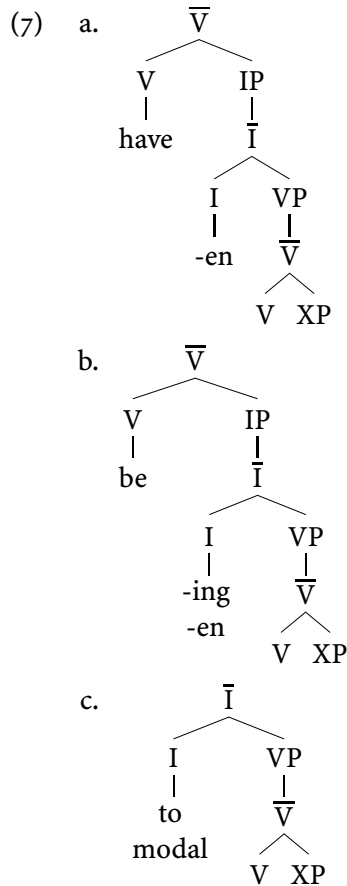


In these parses I've decided to position polarity items as heads that embed the following VP. Moreover, I've renamed these items, and the phrase they head, "Negation." This renaming is common in the literature, as it is typical that only the negative item among the polarity terms is talked about. The representation in which Neg<sup>0</sup> embeds VP is argued for in Pollock (1989), and we will consider his reasons soon. There are other possibilities which are consistent with what we've seen so far, however. One is to treat polarity items as heading a phrase that adjoins, like a modifier would, to the left side of V̄.

This gives a straightforward account of the correlation described above. Verbs stand in the positions determined by the inflections they bear because they are

#### 4. Verb Movement

---



moved to the positions that the phrase structure rules assign to these affixes. The final algorithm determining the surface position of verbs in sentences, then, is a combination of the position that the  $\bar{X}$  Skeleton allows verbs to be generated in, and the action of Verb Movement. Verbs are positioned in a way that is sensitive to both their categorial status and their inflectional status.

An 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 their underlying position. Thus (8) is ungrammatical.

(8) \*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 (9) indicates), and instead English speakers resort to sentences that employ the relatively meaningless auxiliary verb *do*, as in (10).

- (9) \* Gary not likes apples.  
 (10) Gary does not like apples.

Let's for the moment capture this part of the process with the somewhat vague formulation of AFFIX HOPPING in (11).

- (11) 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 have different positions depending on the way in which they are inflected is confirmed by the action of rules that affect VP. These rules seem to necessarily strand an auxiliary verb when it is inflected for agr/tense. Thus, for example, VP deletion and VP topicalization necessarily "strand" a finite auxiliary verb, as (12) shows.

- (12) a. Sam is eating pickles because Mike is  $\Delta$ .  
 b. \* Sam is eating pickles because Mike  $\Delta$ .  
 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.

And, similarly, that main verbs are within the VP they head no matter what their inflectional status is supported by the observation that rules that affect VPs cannot strand main verbs, as (13) indicates.

- (13) a. \* Sam ate pickles because Mike ate  $\Delta$ .  
 b. \* I claimed that Mary put pickles in the toaster, and pickles in the toaster she put.

We have two lines of converging evidence then that auxiliary verbs are spoken in the position our grammar assigns their inflection, whereas main verbs are spoken in the positions they are generated in. This all points to a rule like (14).

- (14) 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 auxiliary verb is necessarily outside of them, it must be that the auxiliary verb cannot remain unmoved.

#### 4. Verb Movement

---

So, in addition to a rule which moves the verb, it is also necessary to explain (or describe) what forces this movement.

One suggestion 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.

(15) 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 to provide the affix in  $I^0$  position with a stem. For main verbs, the STRAY AFFIX FILTER will be satisfied by virtue of Affix Hopping adjoining the bound morpheme onto the following main verb.

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 (16).

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

These are sometimes called Yes/No questions. There is another sort of question which involves a relocation of the finite verb that seeks more specific information. This is sometimes called a Wh-Question, and (17) provides some examples.

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

In both of these types of questions, note, the  $I^0$  surfaces in some sentence-initial position. And as a consequence of this process targeting  $I^0$ , only material which can stand in  $I^0$  can be so affected. Therefore, main verbs in English cannot undergo this process, and instead the auxiliary verb *do* is used.

- (18) 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,<sup>1</sup> is that it is  $C^0$ . The reason for this speculation is that the word-order found 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 (19a) and (19b).

- (19) 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, embedded Yes/No questions 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 questions, then, is (20).

- (20) 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."

- (21) THE WORD CRITERION  
 Let  $\alpha$  be an  $X^0$  immediately dominated by  $\bar{X}$ . Everything  $\alpha$  dominates must form one word.

This will also play a role in preventing word orders like that in (22), which would be created if Verb Movement could adjoin an auxiliary verb to an  $I^0$  occupied by a modal.

- (22) a. \* Gary will have not eaten.  
 b. \* Gary will be not eating.  
 c. \* Gary should have too/so eaten.  
 d. \* Gary should be too/so eating.

Some English speakers may judge (22a,b) grammatical. For these speakers, I suspect that *not* is being used differently than it is in situations where *not* appears between the modal and auxiliary, as in (23).

<sup>1</sup> See den Besten (1983) and Koster (1975).

#### 4. Verb Movement

---

- (23) a. Gary will not have eaten.  
b. Gary will not be eating.

In (23), *not* negates the entire sentence. The speaker of these sentences means to deny the un-negated version of the sentences, i.e. the sentences in (24).

- (24) a. Gary will have eaten.  
b. Gary will be eating.

By contrast, the sentences in (22a) and (22b) are positive assertions. The speaker of these sentences means to assert that Gary should have done something — namely not eaten — or should be doing something — namely not eating. This difference between (22a,b) and (23) can be brought out by considering how they combine with so-called question “tags.” These tags, illustrated in (25), have the opposite polarity as the clause they are appended to.

- (25) a. Gary will eat, won't he?  
b. Gary will not eat, will he?

If the polarity of the tag is the same as the clause it's appended to, the result has a rhetorical flavor:

- (26) a. Gary will eat, will he?  
b. Gary will not eat, won't he?

When *not* stands between the modal and a following auxiliary, adding a positive question tag gives a question of the sort illustrated by (25); cf. (27).

- (27) Gary will not have eaten, will he (have)?

But when *not* follows both modal and auxiliary, adding a positive question tag yields the rhetorical sorts of questions in (26); cf. (28).

- (28) Gary will have not eaten, will he (have)?

On the other hand, when the auxiliary is finite, and so in the  $I^0$  position, a *not* that follows it behaves the same as the one that follows a modal:

- (29) Gary has not eaten, has he?

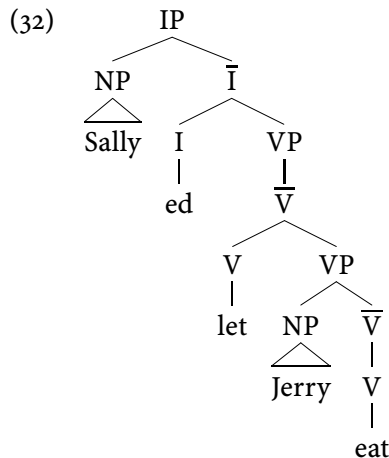
To the extent that there are real judgments here — it's not easy to be clear about what the difference between the “real” and “rhetorical” questions is — it suggests that we do not have the same use of *not* when it stands before and after a non-finite auxiliary. If these uses of *not* correspond to the syntactic position it occupies, then we want to conclude that *not* is in the same position in (30), but a different position in (31).

- (30) a. Gary will not have eaten.  
       b. Gary has not eaten.
- (31) Gary will have not eaten.

This is precisely what the Verb Movement rule, when coupled with the Word Criterion, derives.

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, a sentence whose pre-verb movement parse is that in (32).

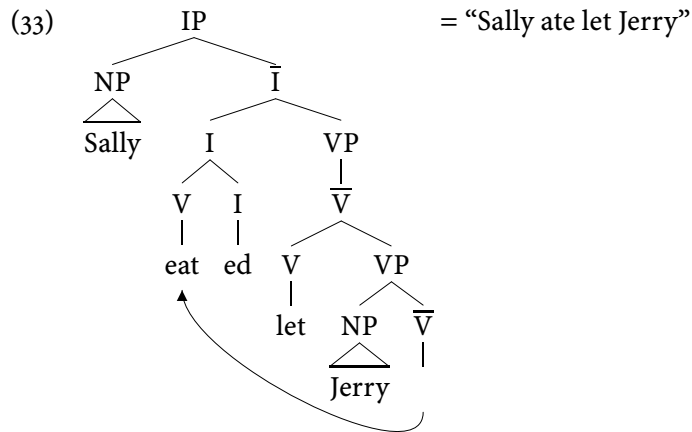


We might expect to be able to get (33) from (32) if *eat* could move past *let* and adjoin to *-ed*, as shown on the next page. Because this doesn't seem possible, it looks like the rule that moves verbs is constrained in such a way that it is unable to move past other verbs.

From the example in (32) 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

#### 4. Verb Movement

---



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 (34).

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

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 (35).

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

There is some independent support for a general constraint of the sort that (35) describes. In the syntax of Noun Incorporation, which arises in some languages when a noun is adjoined to a verb, there are facts which suggest that something like (35) is at play. In Mark Baker's influential study of this phenomena,<sup>2</sup> for example, he argues that (35) 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.

- (36) THE HEAD MOVEMENT CONSTRAINT  
No  $X^0$  may move to a position  $Y^0$ , if there is a  $Z^0$  that c-commands  $X^0$  and is c-commanded by  $Y^0$ .

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.

<sup>2</sup> See Baker (1988).



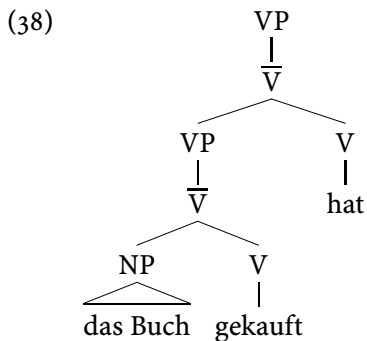
## 4.2 Head Movement's role in "Verb Second" word order

Let's now consider phenomena that look similar in some of the languages closely related to English. We'll 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.

- (37) 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, as indicated in (38).



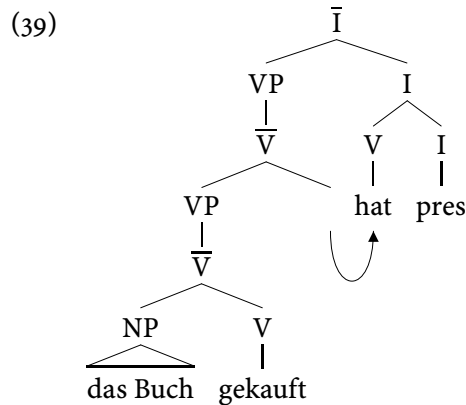
Moreover, if German sentences are IPs, as in English, and finite verbs move overtly to  $I^0$ , then they too must be head final, as shown in (39) on the following page.

Interestingly, we find a different word-order in root, or independent, clauses. Here, the inflected verb no longer comes finally in the series; instead it appears immediately following the subject. So unlike (37), we find only the word orders in (40).

- (40) 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

#### 4. Verb Movement

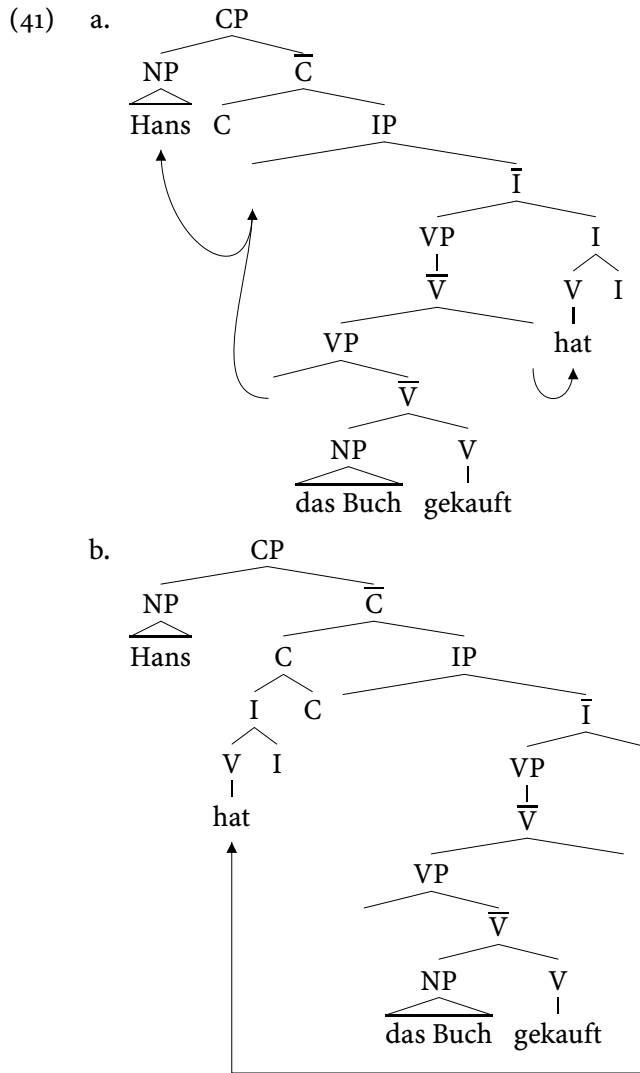
---



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 sits in German immediately follows the subject, and it's into this position that finite verbs are driven by the Stray Affix Filter, or its equivalent in German.

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' positions in root clauses, but something else. Something that distinguishes root from embedded clauses.

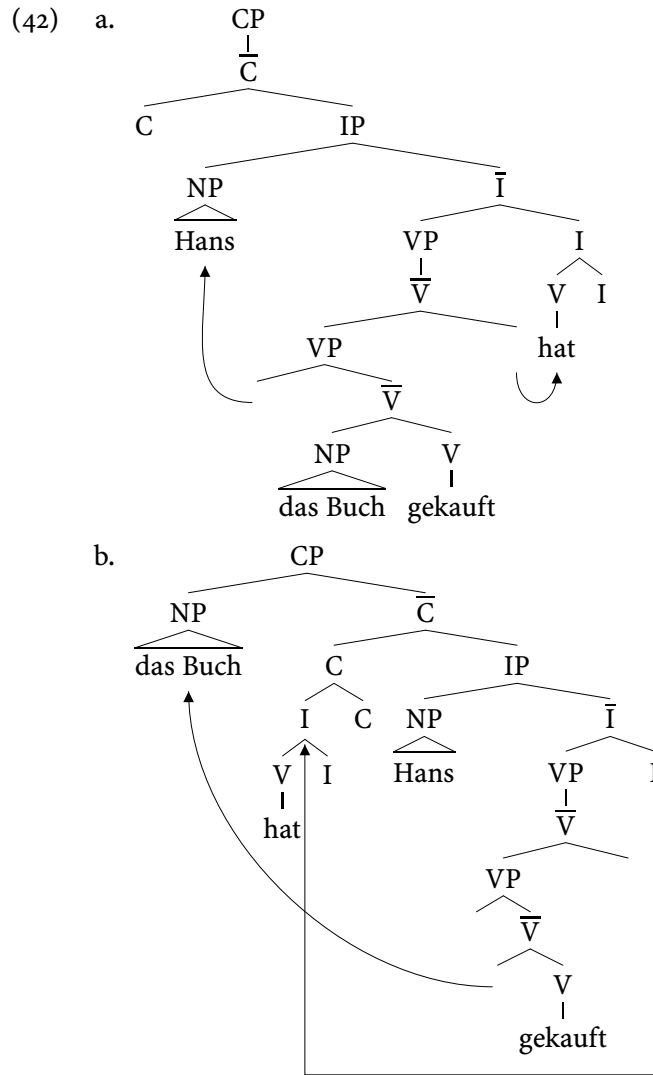
We've already seen a similar difference in the grammar of English: recall that  $I^0$ -to- $C^0$  movement is restricted to root clauses in English. Perhaps the verbs are moving through  $I^0$  into  $C^0$  in cases like (40), then. This would credit German and English with two differences. On the one 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 (40) might get a derivation like that shown in (41) on the next page. Note that in this situation 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 (41) we also find (42) on page 118.



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

#### 4. Verb Movement




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 (43) — and when it does so in German we find V2 word-order, see (44).

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

- (44) 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<sup>0</sup> in main clauses, as in (40a).

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 (45)?

- (45) \* Ich sagte [<sub>CP</sub> das Buch [ daß [<sub>IP</sub> Hans [<sub>VP</sub> gekauft] hat]]].
- 

Vikner (1995) discusses some of the ideas that have been offered for answers to these questions.

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.

- (46) \* 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<sup>0</sup>) 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:

- (47) LIKES ATTRACTS LIKES  
 An X<sup>n</sup> may only adjoin to, or substitute into, a position that is also an X<sup>n</sup>.

#### 4. Verb Movement

---

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 (48).

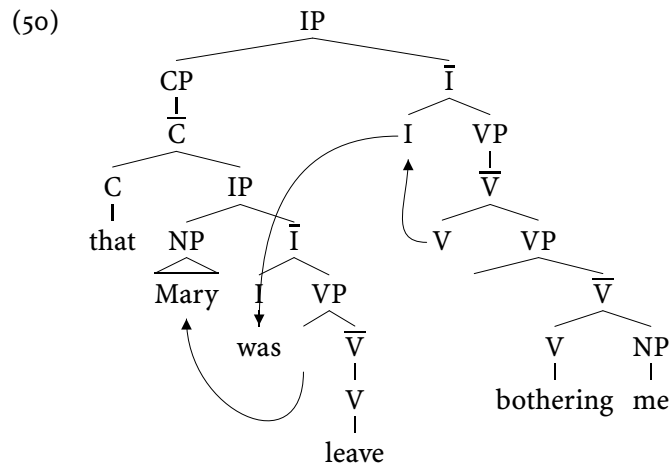
- (48) a. \* ... daß Hans das Buch gehabt kauft.  
       ... that Hans the book had buys  
       b. \* John must had buy the book.

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.

- (49) UPWARDS CONSTRAINT  
 $\alpha$  can move to position  $\beta$  only if  $\beta$  c-commands  $\alpha$ .

That this narrower condition is required for verb movement is shown by cases like (50), which would be possible if verbs could move to non-c-commanding positions.<sup>3</sup>



Let's take a look now at what we find with respect to verb placement in some of the other Germanic languages. In the Scandinavian languages we find a situation similar to German, as Vikner (1995) reviews. Using Danish as an example, note that the same sensitivity to embedding is found in the placement of the finite verb in

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

these languages. As the contrast in (51) 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."

- (51) a. Børnen har set denne film  
 kids-the have seen this film  
 'The kids have seen this film.'
- b. Denne film har børnen set.  
 this film have kids-the seen  
 'The kids have seen this film.'
- c. \* Denne film børnen har set.  
 this film kids-the have seen  
 'The kids have seen this film.'

But, as (52) shows, these traits are not shared by dependent clauses, where, instead, the subject must precede the finite verb.

- (52) 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 (53) show, the placement of the finite verb relative to negation is sensitive to the embedded/non-embedded context.

- (53) 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<sup>0</sup> in independent clauses, just as it does in German.

#### 4. Verb Movement

---

It looks, therefore, like Danish has the word-order of English — placing verbs before their objects — but a 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 (52) (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  $V_2$  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),  $V_2$  word-order is possible even in those embedded clauses which have a complementizer associated with them, as in (54).

- (54) 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 — Vikner (1995) 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, we have the following:

- (55) 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 statements.

- (56) 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<sup>0</sup>?) are movable.
- d. English: auxiliary verbs and I<sup>0</sup> are movable.
- e. German and Danish: Root C<sup>0</sup> and Specifier of CP must be filled. (i.e., V<sub>2</sub>)
- f. English: Only wh-phrases move to Specifier of CP, and only questions trigger I<sup>0</sup>-to-C<sup>0</sup> movement.
- g. Some differences in the nature of I<sup>0</sup>, yet to be explored.
- h. English has affix hopping.

Putting this in the terms of the language acquisition problem which began these lectures, the differences in (56) constitute what the child must learn in order to determine whether the grammar he or she is acquiring is German, Danish or English. We can assume that the other parts belong to the set of Universal Constraints, delimiting the class of possible languages.

But the statements in (56) are not the sorts of things that go into “grammar construction,” in anything like the sense of that process as linguists practice it. Linguists build grammars up from scratch. If the evaluation metric is something like that envisioned in *Aspects of the Theory of Syntax*, then (56) just doesn’t fit. The evaluation metric (56) hints of is not easily characterized in terms of “simplicity,” as in *Aspects*. If (56) 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 preset menu of possibilities just those 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.

### 4.3 The Pollockian revolution: exploded IPs

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, as (57) indicates.

- (57) 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.

#### 4. Verb Movement

---

A word about negation in French is needed before the contrast in (57) will be interpretable. Unlike English, French expresses sentential negation with 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 (57a)). The second word (*pas*) is the one that has a distribution more like that of English *not*. And, as (57) illustrates, the finite form of *have* (*ha*) must precede this part of sentential negation.

French differs from English (but is like German) in moving main verbs. This can be seen in French by the fact that they are placed to the left of sentential negation when they are finite, as (58) demonstrates.

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

The contrast between the position of main verbs in (58a) and (57a) indicates that, as with English auxiliary verbs, 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, which relates to the ordering that the Projection Principle places on arguments and non-arguments. The Projection Principle forces complements to be hierarchically closer to their  $\theta$ -marking head than 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.

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

This apparent difference in the function of the Projection Principle can 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.

Consider next what happens in French non-finite clauses.<sup>4</sup>

<sup>4</sup> These data all from Pollock (1989).

- (60) 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.'
- (61) 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. They may appear either before or after negation, as (62) shows.

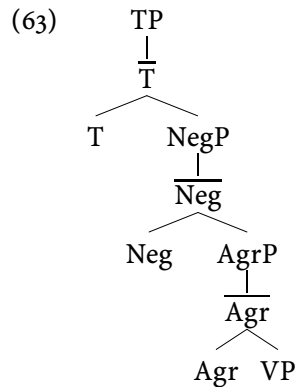
- (62) a. ne *pas* **être** heureux est une condition pour écrire des romans  
 ne not be happy is a condition for writing novels  
 'To not be happy is a condition for writing novels.'
- b. N'**être** *pas* heureux est une condition pour écrire des romans.  
 ne'be not happy is a condition for writing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

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 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.

#### 4. Verb Movement

---

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 (63) 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 pares. 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 (63). It should be noted, however, that this introduces a problem: how is it that  $Agr^0$  can move to  $T^0$  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  $I^0$  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  $I^0$  came in response to the evidence that  $I^0$  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 IndoEuropean languages.<sup>5</sup>

- (64) 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 (65) shows.<sup>6</sup>

- (65) 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."

<sup>5</sup> The argument I'm reporting here comes from Belletti (1990).

<sup>6</sup> These 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).

#### 4. Verb Movement

---

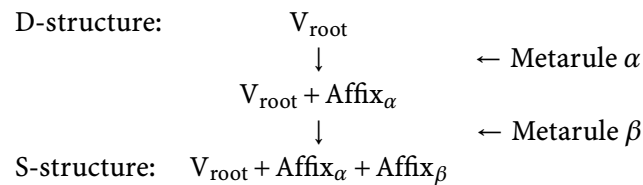
(66) MIRROR PRINCIPLE

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, (66) 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 (67).

(67)

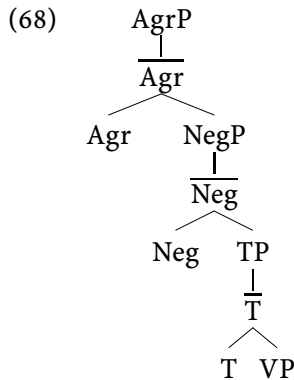


Each metarule brings about the relevant syntactic operation and adds to the verbal root the corresponding affix. Note in particular that this picture correlates syntactic operations with affixes; the correspondences that Baker summarizes with the Mirror Principle only concern the relation between affixal orderings and the orderings of syntactic operations. Indeed, Baker's argument leads to the conclusion that the metarules in (67) 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 (64) 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 that in (68) on the next page.

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:

- (69) 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 (70).

- (70) 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 Neg<sup>0</sup>. Indeed, this correlation holds throughout the Scandinavian languages.<sup>7</sup> 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 (68) predicts.

<sup>7</sup> More 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.

#### 4. Verb Movement

---

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”  $X^0$  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  $X^0$ s 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  $X^0$ s – 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  $T^0$  and  $Agr^0$  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  $X^0$  operation. For English and French, we might assign the values as shown in (71) and (72).

(71) **English:**

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

(72) **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:

- (73) 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:

(74) **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)

(75) **Icelandic:**

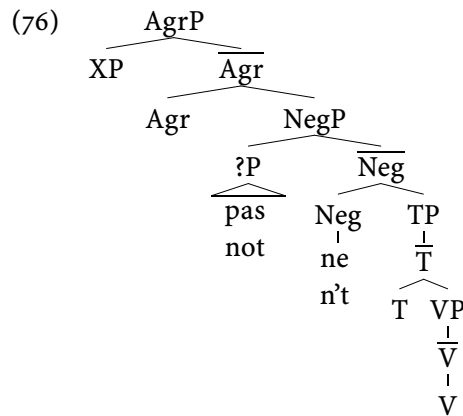
- a. Finite:  
Tense: can't tell  
Agreement: strong
- b. Infinitive:  
haven't looked yet

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  $C^0$ ).

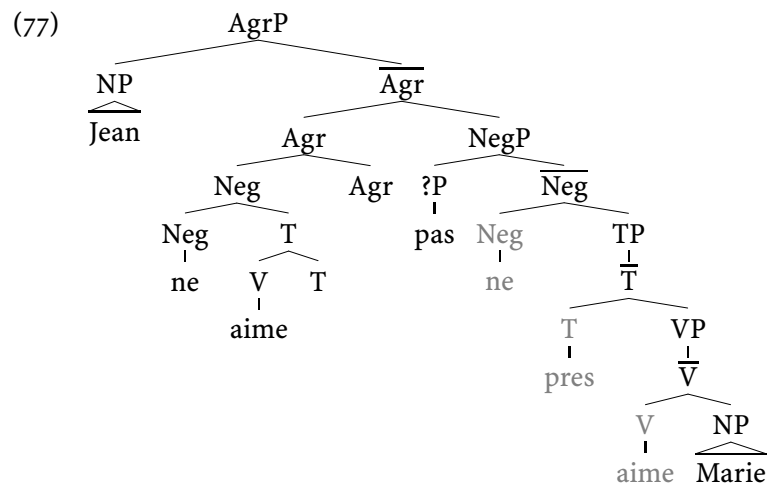
Okay, this is one part of the story. Let's now address the problem that treating negation as a head poses. Recall that we should not expect the Head Movement Constraint to allow movement of  $T^0$  to  $Agr^0$  past negation, if negation is a head that lies between them. Pollock's suggestion is that *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. Thus, the correct way of parsing sentences with sentential negation in them is as in (76) on the following page. On this view, then, when  $T^0$  moves

#### 4. Verb Movement

---



to Agr<sup>0</sup> it doesn't have to skip Neg<sup>0</sup>, it can move through this position and pick up the head (if there is one) on the way.<sup>8</sup> In a finite French clause, then, the surface parse might look roughly as in (77) below.<sup>9</sup> As this parse indicates, *ne* gets carried



<sup>8</sup> This 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.

<sup>9</sup> I will adopt the convention here of putting into a shaded typeface words or phrases in positions they stand in in other parses of the derivation. In (77), for example, the pre-moved position of *aime* is in the head position of VP, and for this reason it shows up in a shaded typeface in that position in this parse.

to the position before *pas* by the tensed verb as it moves into Agr<sup>0</sup>.<sup>10</sup> 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.

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 Agr<sup>0</sup>. In these situations, we might imagine that *ne* moves on its own into Agr<sup>0</sup>. This, in fact, is what Pollock suggests. He hypothesizes that *ne* adjoins to Agr<sup>0</sup> 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 Agr<sup>0</sup>, 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 Agr<sup>0</sup> 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 Agr<sup>0</sup>, we might be better poised to understand why it is linearized to the left of the verb+tense+agreement complex. 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 (78), perhaps.

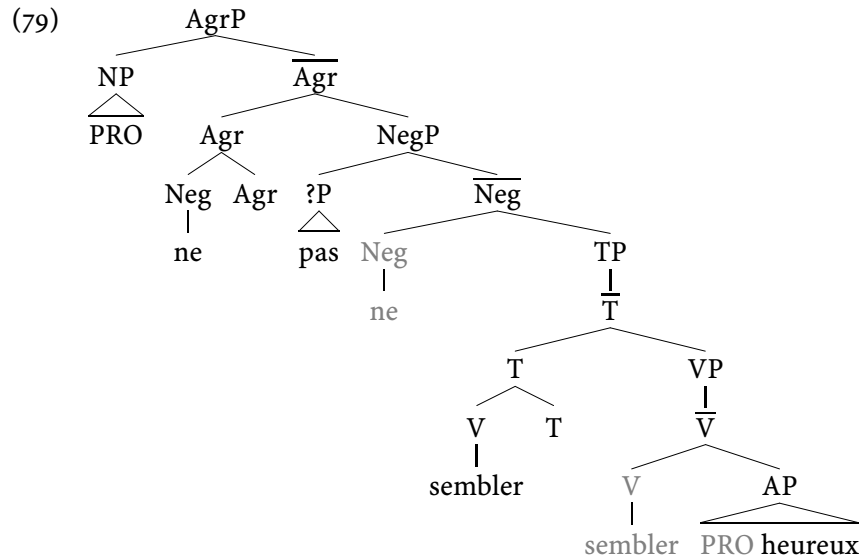
- (78) ADJUNCTION LINEARIZATION  
If  $\alpha$  adjoins to  $\beta$ , then  $\alpha$  precedes  $\beta$ .

With this assumption about *ne*, then, an infinitive that doesn't have verb movement to Agr<sup>0</sup>, but is negated, might have an S-structure like (79) on the following page. So, now what we would like to do is see if we can't get a better understanding

<sup>10</sup> The S-structure in (77) is derived as follows (ignoring the movement of the subject). The main verb, *aime*, is moved into T<sup>0</sup>, where it combines with the tense morpheme. The resulting tensed verb (that is, T<sup>0</sup>) moves and adjoins to *ne*. the resulting negated verb (that is, Neg<sup>0</sup>) moves into Agr<sup>0</sup>, 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.

#### 4. Verb Movement

---



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  $\theta$ -role. In particular, he suggests that we define main verbs as ones that have a  $\theta$ -role to assign, and auxiliary verbs as ones that don’t. He then crafts a proposal that would prevent verbs from assigning  $\theta$ -role 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:

- (80) a. Assume that movement operations leave a “trace” of the moved  $X^0$  or XP. This amounts to treating the shaded copies in my abbreviations of derivations as real terms in the parses that movement operations create, and calling them “traces.”
- b. Let the Theta Criterion be enforced at S-structure.
- c. Let the trace of a verb be capable of assigning the verb’s  $\theta$ -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 (82), nor for the contrast in (81).

- (81) Mary hasn’t a dime.

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

Why doesn't the *be* in (81a) have the same number of  $\theta$ -role as (81b)? And it sure looks like *has* in (82) is capable of moving past negation, though it appears equally like it assigns  $\theta$ -role. Pollock suggests that the  $\theta$ -role assigned in (82) comes not from the verb, but from some hidden term. Something similar might be attempted in distinguishing the cases in (81).

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; (83) is an example, for instance, where the main verb has relocated to  $C^0$  position.

- (83) 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 (71), (72), (74) and (75). 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.

## 4. Verb Movement

---

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  $\text{Agr}^0$  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 (84).

- (84) A main verb adjoined to  $\alpha$  at S-structure creates ungrammaticality unless there is a “strong” morpheme in  $\alpha$ .<sup>11</sup>

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.


### 4.4 Features and covert movement

Let’s take a closer look at Affix Hopping, now, and try to devolve this operation into the schema of Move  $X^0$  and morpheme strength 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 (85) and (86).

- (85) a. Jerry -s make [<sub>VP</sub> Sally eat].  
b. \* Jerry make [<sub>VP</sub> Sally eats].

<sup>11</sup> This 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  $C^0$ .

- (86) a. Jerry should remember [<sub>IP</sub> Sally -ing eat].  
 b. \* Jerry should remembering [<sub>IP</sub> Sally eat].
- 

In (85), 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.

In (86), 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, (87) seems to be true:

- (87) When an affix,  $\alpha$ , joins with a verb,  $\beta$ , by way of Affix Hopping, it must be that  $\beta$  could have undergone Move  $X^0$  to  $\alpha$ .

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), (87) spreads the inverse of these constraints onto Affix Hopping. As (87) 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 that can be expressed with (88).

#### 4. Verb Movement

---

(88) If Move  $X^0$  can bring  $\alpha$  and  $\beta$  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:

(89) \*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.<sup>12</sup> His proposal doesn't capture all of the properties of Affix Hopping that we need to explain — it leaves out what makes (89) bad, for example — but it does point in a direction that has been influential, and that does a good job of explaining (87) and (88). 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 (88) 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:

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

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 (90) would be a way of explaining the similarities that

<sup>12</sup> What I will present here is a pared down version of his proposals in chapter 3 of Chomsky (1995b).



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.

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 use the Word Criterion/Stray Affix Filter to force a verb to move into a position in which a bound morpheme stands. Instead we're going to have to formulate a condition that requires a head with a tense or agreement feature to be paired with a verb bearing tense or agreement morphology. Chomsky suggests doing this with the following two principles:

- (91) An inflectional feature must be "checked."
- (92) An inflectional feature is checked when it shares an  $X^0$  with a term that bears matching morphology.<sup>13</sup>

Like the Stray Affix Filter, or Word Criterion, 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 checked.

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

<sup>13</sup> A 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" born by  $T^0$ , and so on.

## 4. Verb Movement

---

in (90). 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  $\theta$ -role from and the position it gets its Case from are often different, so also are movement transformations used in solving these similar dilemmas.

Our purpose here is different with respect to the semantic outcome, however. In those cases, the theory is designed to move NPs covertly into positions in which they seem to be semantically interpreted. In the cases at hand, we want to move verbs covertly into the positions that their inflections are related to, but we want to interpret them semantically in their pre-moved positions. We would not want verbs to gain arguments in the positions they move to, for instance. There is no semantic reason, then, for believing that the way verbs covertly move should be related to the way that we will see NPs covertly move.

But this is all a bit premature. We have yet to see the covert movement of NPs that is being alluded to here. I will therefore describe the technique Chomsky devises in a way that does not make reference to this other process.

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 (or more) of which is spoken, the s-structure(s). These sets of parses are formed by taking a d-structure and generating a collection of other parses by applying rules such as Argument Movement, NP Shift and Head Movement. Then these rules apply again to each of the members of the derivation created by the first pass. And then these rules again apply to all the members of the set created by the second pass. And this process repeats until no more additional parses can be formed.<sup>14</sup> There has been

<sup>14</sup> For the simple cases we’ve examined we could do with a simpler way of generating derivations. We could let the rules operate on the d-structure and then stop. How these two processes differ will

no need up to now to give any additional structure to this collection of parses. It has been sufficient to think of a derivation as a simple set. The task ahead of us requires, however, that we have sufficient control over this collection of parses to be able to manipulate which of these are s-structures. One way of gaining this control is to order these parses into a series. This is what Chomsky's proposal requires.

Therefore, let's order the parses in the following way. The first parse will be the one from which all the others are generated: the d-structure. Every other member of the derivation is derived from the immediately preceding one by application of a single rule.

- (93) Let  $\mathcal{P} = (P_1, P_2, P_3, \dots, P_n)$  be a derivation for some sentence, S, if:
- a.  $P_1$  is the d-structure for S, and
  - b. Each  $P_{i+1}$  is derived from  $P_i$  by the application of exactly one rule.

Note that this does not define a unique derivation for any given sentence. Because there are no constraints on the order in which the rules must apply, there are quite a number of different derivations for any given d-structure. Each sentence, then, is associated with a (sometimes large) set of derivations. I haven't done the proof, but I believe the parses that exist in this set of derivations will end up being the same as the collection of parses that are found in the previous, less structured, definition of derivation. All that has been added here, then, is some structure to the set.

Let's examine now how the s-structure representations in these derivations are determined.

Consider first those derivations in which only Argument Movement is used. It turns out that for all these derivations,  $P_n$ , the final member, will be the sole parse that meets the requirements on s-structures. This emerges as a consequence of how Argument Movement is constrained. The first relevant constraint is the one that allows Argument Movement to relocate an argument only to a c-commanding position. This constraint means that there will be no derivation in which a phrase in  $P_{i+1}$  is in a "lower" position than it is in  $P_i$ . For all phrases that Argument Movement relocates, then,  $P_n$  will be the parse in which they have their "highest" position. Consider now an application of Argument Movement to satisfy the 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,  $P_n$ , and no other parse, of every derivation created by Argument Movement alone will have all overt NPs in Case marked positions. Because we have defined s-structures as those parses that satisfy the Case Filter,  $P_n$  will be the s-structure for all of these derivations. Similarly, if Argument Movement is invoked to create a parse that satisfies the condition

depend in large part on how the rules are formulated and constrained.

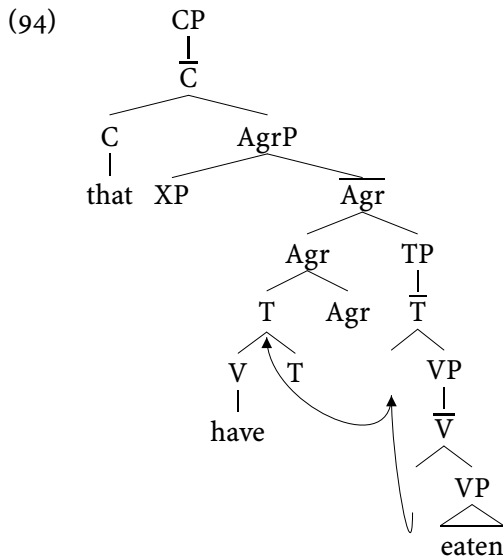
#### 4. Verb Movement

---

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 – this, recall, is the requirement on the s-structure parse containing a PRO. Argument Movement 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. As we've seen, PRO is allowed in the Specifier positions of certain phrases. It can be in the Specifier position of an infinitival complement CP, for instance. But because Argument Movement is prevented from moving things out of CPs, PRO will never be able to move beyond this position. PRO is also allowed in the Specifiers of phrases in adjunct or Specifier position, but there are constraints on movement we have not yet encountered that will prevent movement from phrases in adjunct of Specifier position. As it happens, then, the positions in which PRO is allowed are also positions from which Argument Movement cannot move them further. As a consequence, for all derivations that are created just by Argument Movement,  $P_n$  will be the sole parse that satisfies the well-formedness conditions on s-structures that contain PRO.

Consider now those derivations which Argument Movement and Heavy NP Shift together generate from some d-structure. Consider the first parse in these derivations in which the Case Filter and the PRO Restriction are both satisfied. This will be a parse that qualifies as an s-structure. As we've just seen, there can be no parse following this one that is generated by Argument Movement. So, if there are any subsequent parses, they will all be generated by Heavy NP Shift. So far as we've seen, all these parses are pronounceable. All of them qualify as s-structures. For this class of derivations, then, there will be some series of contiguous parses which includes  $P_n$ , all of which are s-structures.

Finally, consider the class of derivations that are generated from some d-structure by application of Argument Movement, Heavy NP Shift and Head Movement. These derivations will have the same properties as those derived by just Argument Movement and Heavy NP Shift, except with respect to the placement of the moveable heads. Suppose that we take the requirement that inflectional features be checked (i.e., (91)) to be a well-formedness requirement on s-structure. For the cases we've examined, this will have the effect of preserving the earlier outcome that the s-structures will be some series of contiguous parses that includes  $P_n$ . This is because in all the cases we've looked at, the highest  $X^0$  with a feature in need of checking has as its closest c-commanding head something that the Word Criterion would prevent head movement into. For instance, in a derivation in which a finite auxiliary moves to  $T^0$  and  $Agr^0$  to check the features these heads come with, the first parse in the derivation in which these features are checked is something like (94).



There are no later parses in this derivation formed by Head Movement. The Head Movement Constraint prevents  $\text{Agr}^0$  from moving beyond  $\text{C}^0$ , and the Word Criterion prevents  $\text{Agr}^0$  from moving to  $\text{C}^0$ .

For all the cases we've examined, then, the s-structures will always be found at the end of the derivations, and include the final parse. The method of "covert movement" that Chomsky invokes in these contexts is to define S-structure in such a way that it does not necessarily include 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. Because what we want to achieve here is covert Head Movement, this can be done by taking the requirement features get checked to be satisfied even if it doesn't hold of the spoken parse. That is, take it to be a condition like the EPP that is satisfied if the derivation has one parse that satisfies it. This will permit derivations in which the spoken parse precedes an application of Head Movement. It will allow verbs, and other heads, to be spoken in their pre-moved positions.

If we adopt a view of derivations like this one, and hold constant the view that the PRO Restriction and the Case Filter are well-formedness conditions on s-structures, then we have (95).

- (95) 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.

#### 4. Verb Movement

---

- c. Let *s-structure* be a phrase marker that satisfies the Case Filter and the PRO restriction and is phonologically interpreted.
- d. A *Derivation* =<sub>def</sub> an ordered n-tuple of phrase markers,  $\mathcal{D} = (P_1, P_2, \dots, P_n)$ , such that:
  - i. Each  $P_i = \mathcal{R}(P_{i-1})$ , and
  - ii. The first member is a d-structure, and
  - iii. There is at least one member that is an s-structure, and
  - iv. There is some  $P_i$  in which all inflectional features are checked.
  - v. For every IP,  $\alpha$ , in the d-structure, there is some phrase marker which has a phrase in Specifier of  $\alpha$ .<sup>15</sup>

(96) A grammatical sentence must have a well-formed derivation.

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:

(97) An s-structure must have all strong features checked.<sup>16</sup>

This can be put into the terms (95) uses to define derivations with (98).

(98) Let an *s-structure* be a phrase marker that satisfies the Case Filter, the PRO restriction, and **has no unchecked strong features in it** and is phonologically interpreted.

What (98) does is, essentially, enforce the Word Criterion/Stray Affix Filter at S-structure for those  $X^0$ 's that have a strong feature. Thus, (98) 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  $\text{Agr}^0$  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 (98), then, is that

<sup>15</sup> This is the Extension to the Projection Principle.

<sup>16</sup> Chomsky actually proposes that unchecked 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.

the only time covert Verb Movement can be used to satisfy (91) is when an  $X^0$  has a weak feature.

Therefore (98) correctly forces overt verb movement in cases where strong features are found. It allows covert Verb Movement to occur only when a verb is checking off a weak feature. Now, what we have to do is find a way of making sure that when this verb is a main verb, not only is covert movement permitted, it is forced. Chomsky suggests we do this with a principle that favors covertly moving terms to overtly moving them. Of the various ways of expressing this idea, one that is found in the literature is (99).

(99) EARLINESS

Let  $\mathcal{D} = \{D_1, D_2, \dots, D_n\}$  be the set of well-formed derivations for some sentence,  $S$ , and  $\mathcal{O}$  be the set of d's, such that for every  $D_i \in \mathcal{D}$ ,  $d_i$  is that sub-series of parses in  $D_i$  that starts with the d-structure of  $D_i$  and ends with the first s-structure in  $D_i$ . The s-structure of  $S$  is the one in the shortest  $d \in \mathcal{O}$ .

What Earliness does is find the derivation that produces an s-structure with the minimal number of steps, and selects this as the s-structure for the sentence. As a consequence, 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 Earliness. In other words, when main verbs check off weak features, they must do so in the post-s-structure section of the derivation.

An interesting property of EARLINESS is that it removes the accidental nature of the fact that derivations that are produced by A Movement alone always terminate in an s-structure.

The last part of the system that we must consider arises in cases where an auxiliary verb is responsible for checking inflectional features. In English, we want Head Movement to feed s-structure in this scenario: when  $\text{Agr}^0$  and  $T^0$  have features, as they do in finite clauses, those features are checked by way of overt verb movement. Thus we need a way of forcing auxiliary verbs to move to check the weak features of  $\text{Agr}^0$  and  $T^0$  in English. Chomsky does this by stipulating that auxiliary verbs cannot move covertly. I will explore another avenue, one that's more in keeping with the feature-based system we are examining.<sup>17</sup> Let us imagine that the finite morphemes that auxiliary verbs bear have features. This will make auxiliary verbs move to a matching  $X^0$ , no matter whether that  $X^0$  has features. If we imbue the features on the auxiliary's inflection with the strong flavor, this will make this movement feed s-structure. Thus, I suggest:

<sup>17</sup> I owe this suggestion to Angelika Kratzer.

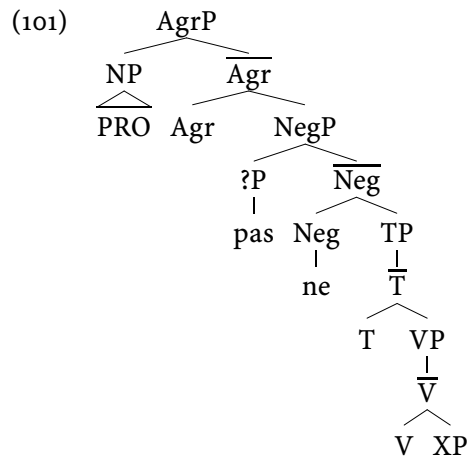
#### 4. Verb Movement

---

(100) The agreement and tense morphology on auxiliary verbs have a strong feature.

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 (101) below before Head Movement has applied. What we desire is to let the verb move to  $T^0$ ,



and to allow auxiliary verbs, 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.<sup>18</sup> 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  $T^0$ .

<sup>18</sup> Recall that the reason Pollock gives 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. 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



Now, what about Agr<sup>0</sup>? In my earlier characterization of French infinitival Agr<sup>0</sup>, I suggested that it was empty. If we continued to characterize it this way, then Agr<sup>0</sup> would presumably lack features altogether; and this, of course, would mean that it does not have strong features. As a consequence, (98) will not force verbs to overtly move to Agr<sup>0</sup> in French infinitives. When verbs aren't forced to move overtly, then Earliness is going to force them to move covertly. That means that neither main nor auxiliary verbs should be able to overtly move to infinitival Agr<sup>0</sup> in French. That's not the result we want, because, as we've just reviewed, it's only main verbs that cannot move to Agr<sup>0</sup> 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 T<sup>0</sup>, and moves, optionally, into Agr<sup>0</sup>. But as before, what can the features on Agr<sup>0</sup> be that will allow optional overt movement of *to*? If they are strong features, then overt movement will be forced, and if they are weak features, then Earliness will force movement to be covert.

In fact, as these two cases make clear, one of the consequences of Earliness 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 this proposal, then. There is no solution to these problems in the literature (that I am aware of). We're on our own. Let me make the following suggestions. For English, let us adopt the hypothesis that *to* can be generated in either T<sup>0</sup> or Agr<sup>0</sup>. This amounts to the claim that we have two *tos* in English, one that is an Agr<sup>0</sup> and another that is a T<sup>0</sup>. On this view, we will not have to move *to* into Agr<sup>0</sup> from T<sup>0</sup>, and so there is no consequent optional movement. For French, I suggest that we let Agr<sup>0</sup> have weak features but give to the infinitival morphology on auxiliary verbs a feature, much as we did for English auxiliary verbs. Just as for the finite morphology on auxiliary verbs in

an adverb *to* fall 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<sup>0</sup>.

#### 4. Verb Movement

---

English, let's let the infinitival morphology on French auxiliary verbs be strong. This will force auxiliary verbs to move overtly to Agr<sup>0</sup> in order to check this feature that they carry. To make this movement optional, we must let this feature only optionally be present on the auxiliary verb's infinitival morphology. I don't find this a satisfying, or informative, proposal; but I can think of no other.

This isn't the only context in which we'll encounter difficulty for the view that overt movement is obligatory, however, and so we might want to seek a more general solution. Note, for instance, that EARLINESS disallows Heavy NP Shift. In our original model, it was possible to say that a sentence was associated with more than one s-structure, and, as we've seen in our discussion of serialized derivations, these additional parses are added by Heavy NP Shift. Because EARLINESS requires that we select just one s-structure, however, it will no longer be possible to model the Heavy NP Shift word orders in this way. We'll have to revisit this issue, then. But let's do so once we've had a chance to investigate more thoroughly the constraints on Heavy NP Shift (something we will do in chapter 6).

Until then, 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 inflectional morphology are brought together by Head Movement, it is no longer equipped to express Baker's Mirror Principle.

# 5

## 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 of non-trivial strings of words 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, as (2) illustrates).

- (2) a two 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

## 5. Determiner Phrases and Noun Movement

---

transparent, something along the lines of (3), where the ~~struck-out~~ material should be understood as syntactically present, but phonetically absent.<sup>1</sup>

- (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 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; that is why examples such as (??c) are ungrammatical.

- (4) a. Mary's lamp  
b. the lamp  
c. \* the Mary's lamp

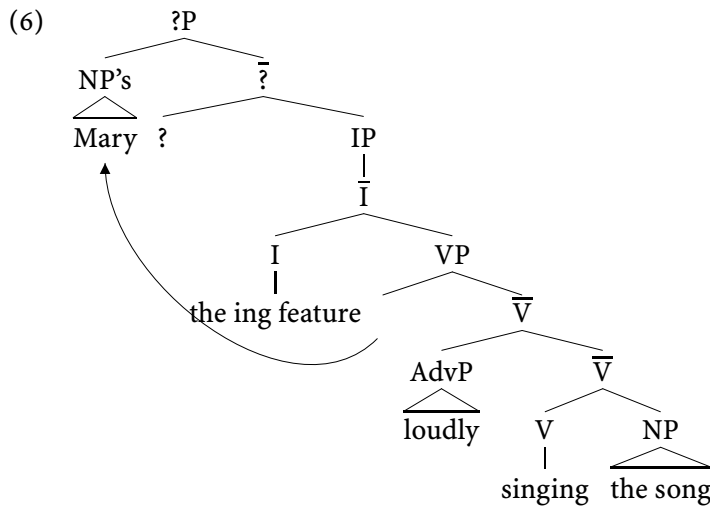
We adopted a view of NPs that embraced a constraint — yet to be found — that limited DPs and genitive NPs to their Specifier position. That is, this fact was encoded in the Phrase Structure rules we began with in the way that NPs are expanded. But when we transitioned to a category-free set of phrase 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.

In figuring out what principles are responsible for positioning possessives and determiners in the same, initial, position, we will want to find what is responsible for fixing the positions of these rather different terms. To capture the fact that they compete for the first position in an NP, we should design these principles so that they allow them only this spot. In fact, however, we'll find that there are some places that possessives can be in that don't look at all like NPs; (5) are some examples.

- (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.

<sup>1</sup> For an examination of cases like (3b), see Hackl (2000).

These strings fit the description of a clause, or VP, except that they begin with genitive nouns. This suggests that these phrases have an organization something like (6).



But what's the “?” in this graph?

### 5.1 The DP Hypothesis

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. This is also true for the expressions in (5):

- (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.”)

Moreover, the expressions in (5) can be conjoined with ‘normal’ NPs, which, if we've got the constraints on Coördination correct, also indicates that they are NPs.

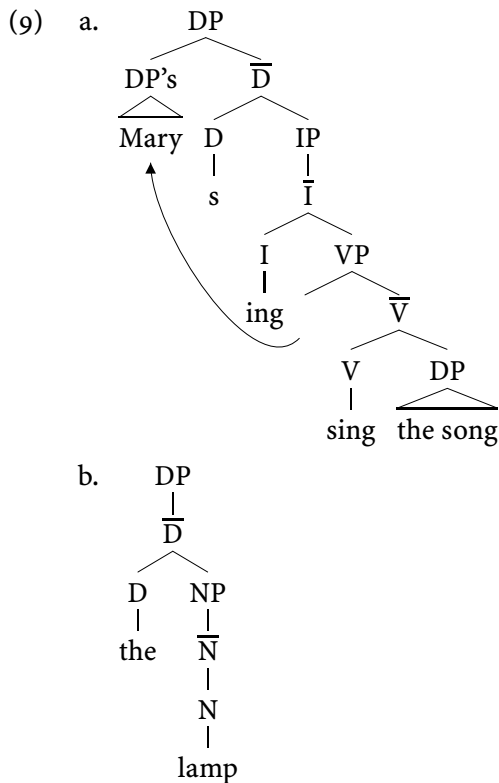
- (8) [Mary's singing the song] and [my subsequent departure] enraged the organizers.  
 (compare: “\*[Mary's singing the song] and [that I subsequently departed] enraged the organizers.”)

But if ?P is a noun phrase, then the law of endocentricity demands that ? be a noun, contrary to fact. Something's amiss.

## 5. Determiner Phrases and Noun Movement

---

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.<sup>2</sup> 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

<sup>2</sup> I've changed slightly his discussion of cases like “Mary's singing the song” — but the spirit of the argument is his.

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  $A^0$ , 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 expressions like *everyone/everything* and *someone/something*:

- (12) a. someone angry  
 b. something large  
 c. \* angry someone  
 d. \* large something
- (13) a. everyone angry  
 b. everything large  
 c. \* angry everyone  
 d. \* large everything

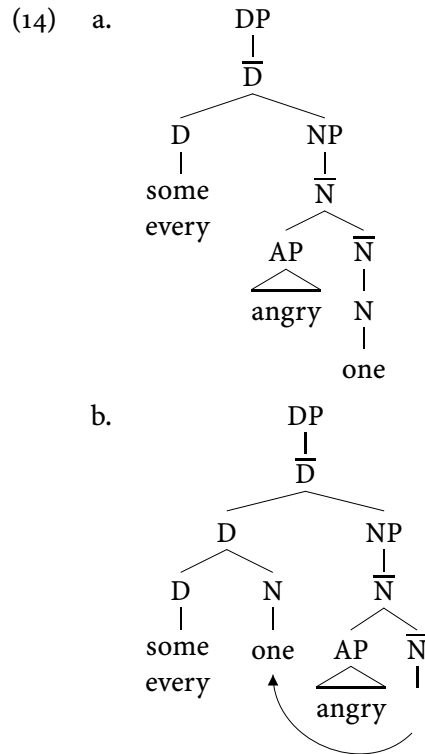
Abney<sup>3</sup> suggests that an analysis of this exception should not make it accidental that the determiners *every* and *some*, and the noun *one* and *thing* are involved. More particularly, it should not be accidental that the only expressions in English which seem to be made up of a determiner 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/thing* onto the determiner; that is, he suggests that (12) be involved in a derivation that includes the parses in (14) on the following page. 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.

<sup>3</sup> Who is here following a suggestion of Richard Kayne's, who in turn is building on ideas in Postal (1969).

## 5. Determiner Phrases and Noun Movement

---



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.<sup>4</sup>

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

<sup>4</sup> We 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).



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.<sup>5</sup>

In this chapter, we will examine a paradigm of facts which focus on the noun movement part of this parallelism.

## 5.2 Noun Movement

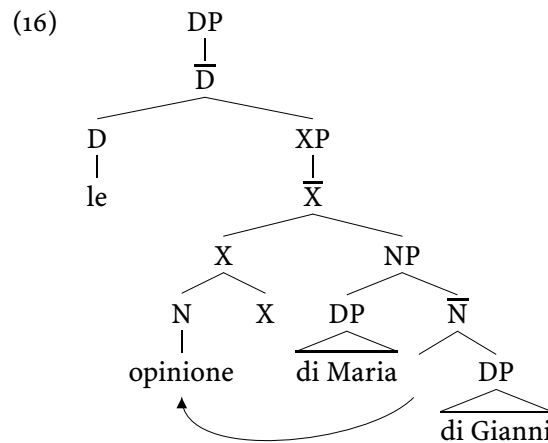
One paradigm of facts that has been accounted for in terms of noun movement concerns a difference in the position of a noun's "Subject," which is how we might characterize the terms that appear as genitives in English. In Romance (by which I will mean here Catalan, standard French and Italian), the "subject" argument can appear between the noun and its complements.

- (15) a. L'opinione di Maria di Gianni (Italian)  
 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 "overt"  $N^0$  movement in Romance but not English. If we understand the Derived Subjects Hypothesis to claim that it is the highest  $\bar{N}$  which assigns the  $\theta$ -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. This derivation is outlined in (16), which corresponds to the Italian example in (15a).

<sup>5</sup> The first two chapters in Abney (1987).

## 5. Determiner Phrases and Noun Movement



I think the first argument of this sort comes from Cinque,<sup>6</sup> who makes the argument based on the position of “ethnic” adjectives, which also can be found post-nominally in Romance.

- (17) L'invasione tedesca dell'Austria.  
the invasion german of Austria

Ethnic adjectives seem able to bear a subject  $\theta$ -role assigned by a noun. So, consider, for example, the contrast in (18) below.

- (18) 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 (18a), *American* serves as a genuine modifier, merely attributing to the referent involved that it has the property of being American. But in (18b), *American* refers to an abstract entity that is constituted of the American people or the American government.<sup>7</sup> That this is a function of these adjectives bearing the “subject”  $\theta$ -role which *opinion* and *discussion* assign is indicated by the fact that this meaning is lost when there is another external  $\theta$ -role bearer in the nominal.

- (19) a. Uganda's American opinion of the blockade  
b. Morocco's American discussion of trade barriers

<sup>6</sup> In an unpublished talk delivered at the Università di Venezia in 1990. See his GLOW abstract from the 1992 Lisbon meeting.

<sup>7</sup> This quality of ethnic adjectives is first discussed, to my knowledge, in Kayne (1984a, Chapter 7).

Like other external  $\theta$ -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 *di/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.

- (20) \*L’invazione dell’Austria tedesca  
       the invasion of Austria   german  
       ‘the Austrian invasion of Germany’                   (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 (21), compare (22).

- (21) a. Iran’s bombardment by Russia took several weeks.  
       b. Uganda’s invasion by Tanzania grinds slowly on.  
 (22) 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 (21) have undergone A Movement from a position to the right of the noun, where they receive their  $\theta$ -role. Ethnic adjectives, apparently, are unable to move from this position. Instead, they are stuck in the position from which they get their  $\theta$ -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  $\theta$ -role is assigned

## 5. Determiner Phrases and Noun Movement

---

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 they move to a position that is associated with the number morphology on the noun. 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.

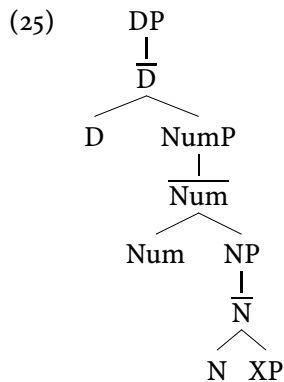
- (23) a. ea rea kaaroo neey  
sing car this  
b. ea gal kaaroo neey  
dual car this  
c. ea pi kaaroo neey  
plural car this

This at least suggests that Universal Grammar makes projecting a syntactic phrase above Number a possibility. Further, Dryer shows that the relative order of Num<sup>0</sup> and Noun correlates with Verb-Object word order. This would be explained, on standard theories of word order typology, if Num<sup>0</sup> 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 still above adjectives and the noun. He provides examples like the following.

- (24) 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 (25) on the facing page. This is what the statistical study yields.

Of course, if Dryer's conclusion that adjectives come between Num<sup>0</sup> and N<sup>0</sup> is valid for English, then English nouns must combine with this position 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, which is even higher. Or, alternatively, we might imagine that the noun undergoes covert movement to  $\text{Num}^0$ .

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  $\text{N}^0$  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  $\text{Num}^0$ . 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.

- (26) 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.

## 5. Determiner Phrases and Noun Movement

---

- (27) 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 (27a), the nominal refers to an individual drawn from the set of things that are large valleys. In (27b), 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 (28a) the PP *from Duluth* stands in the same relation to *Mary* as it does in (28b).

- (28) a. Mary, from Duluth, has arrived.  
b. Mary has arrived and she is from Duluth.

In an example such as (29), by contrast, *from Duluth* plays a more direct role in determining the reference of the DP it is attached to.

- (29) 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 (29) with a circumlocution, parallel to (28b), like:

- (30) 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.

- (31) 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.

- (32) \*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 (32) 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.

- (33) 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.

## 5. Determiner Phrases and Noun Movement

---

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.

- (34) a. dès r'tchâfés cromptîre[s]  
the reheated potatoes (Walloon)
- b. des petites fille[s]  
the small girls (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.

- (35) a. un mal  
an evil
- b. des maux  
evil(s)
- (36) a. un oeil rouge  
a red eye
- b. des yeux rouges  
red eye(s)
- (37) a. quel cheval  
which horse
- b. quels chevaux  
which horses

In Walloon, however, these forms always appear just in the same, uninflected, form. Examples are in (38)-(40).

- (38) a. on mâ  
evil
- b. dès mâ[s]  
evils
- (39) a. on rothch[e]-oûty  
red eye



- b. dè̃s rodj[e]-z-où̃y  
red eyes
- (40) a. [ke: dzva:]  
which horse
- b. [ke: dzva:]  
which horses

Further, she notes that liaison is possible in literary French, but never in Walloon.<sup>8</sup> Liaison refers to a phonological process that allows the normally not pronounced plural suffix to be pronounced in literary French in certain contexts. Typically these contexts are ones in which the word following the plural suffix starts with a vowel, like the following.

- (41) Les train-z-ont du retard.  
the train-s-are running late

In this example, the plural suffix on the end of *train* is pronounced because the following verb starts with a vowel. To account for why this process is available in French, it would seem necessary to imagine that the plural suffix is, in some sense, present on French nouns, even if it is normally not pronounced. Its absence in liaison contexts in Walloon, then, could be blamed on its total absence in that language.

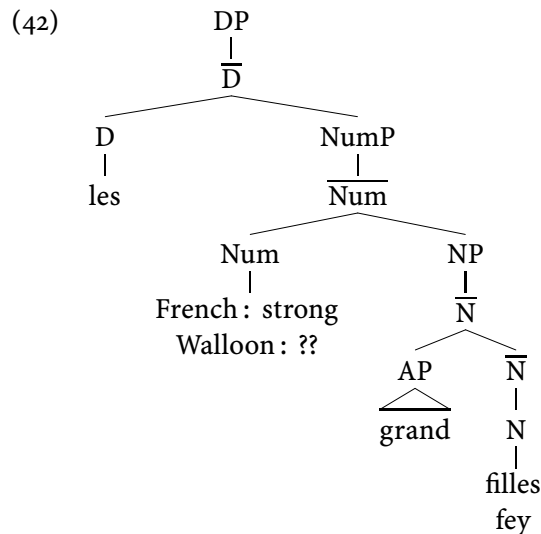
From these data, then, she concludes that the plural affix in French is *-es*, or a suppletive trigger, and that it is absent altogether in Walloon (well, almost — we'll revise this immediately). Thus, using the inflectional feature model we adapted from Chomsky in connection with verbal syntax, this gives to French/Walloon nominals the d-structure in (42) on the next page. In French, the head noun is driven overtly into the Num<sup>0</sup> 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 nominal forms in French is correlated with these noun's 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

<sup>8</sup> But 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).

## 5. Determiner Phrases and Noun Movement

---



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 Num<sup>0</sup> 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 Morin (1986), that leads her in this direction is that liaison between prenominal adjectives and a following noun is absent in Walloon, though present in French. The normally silent [z] ending *gros* and [t] ending *petit* are pronounced with a vowel initial noun follows in French, as in (43).

- (43) a. un gro-z -arbre  
a big tree
- b. une peti-t -enfant  
a little child

- (44) a. on gro[s] abe  
a big tree
- b. on peti[t] èfant  
a small child

But this doesn't happen in parallel circumstances in Walloon, as in (44).<sup>9</sup> Bernstein suggests that her account would provide an immediate explanation for this, if in Walloon there is a Num<sup>0</sup> 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 Num<sup>0</sup>, and is therefore in a position to trigger liaison with a preceding adjective. For this reason, she suggests that Num<sup>0</sup> should be present in Walloon, but not able to attract N<sup>0</sup>s to it.

Note how this account presupposes that adjectives cannot come between Num<sup>0</sup> and N<sup>0</sup>. If they could, then an intervening Num<sup>0</sup> 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 Num<sup>0</sup>.<sup>10</sup> 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. It should also be noted that this would diverge from the trend Dryer found in the relative placement of number words and adjectives. His results suggest that adjectives should be placed lower in the DP than number Num<sup>0</sup>.

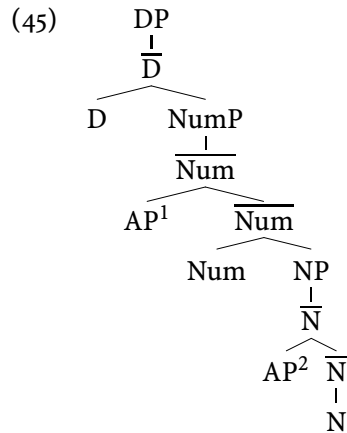
In fact, we have another kind of problem that arises as soon as we adopt the view that it is movement of Nouns to Num<sup>0</sup> that is responsible for the N+Adj word order. In the system we have adopted from Chomsky, there is no way of making head movement systematically optional. And, as we have 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 Num<sup>0</sup>. One way we could do this is to imagine that adjectives can be placed either above or below Num<sup>0</sup>, as indicated in (45) on the following page. Once the noun has moved into Num<sup>0</sup>, there is still, according to this model, a place for adjectives to the left of the Num<sup>0</sup>+N<sup>0</sup> pair. In order to account for the meaning difference that (sometimes) arises, we might imagine that adjectives in the AP<sup>2</sup> position get a different interpretation (maybe restrictive) than do adjectives in the AP<sup>1</sup> position. We might seek an account for this difference from the fact that these adjectives are modifying different things: a Num<sup>0</sup> in one case and a N<sup>0</sup> 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 ad-

<sup>9</sup> (44a) is Liège Walloon and (44b) is from the Goundecourt Picard dialect.

<sup>10</sup> It 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.

## 5. Determiner Phrases and Noun Movement

---



jective 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.

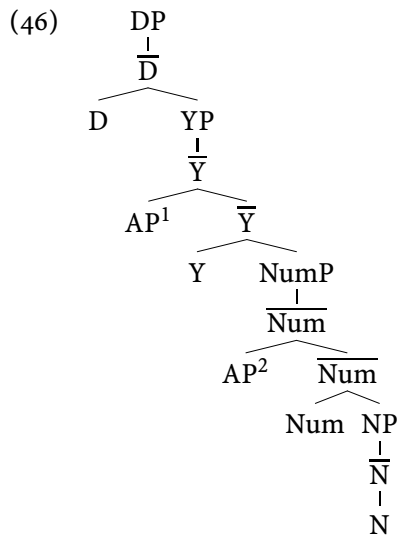
A different way of modeling the N+Adj/Adj+N word order, that still correlates the availability of the N+Adj order with overt Noun Movement to Num<sup>0</sup> and would also address the problem we encountered with Bernstein's explanation for the contrast in liaison between Walloon and French, would be 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 in "Y" here. The idea, then, would be to give to DPs the shape in (47) on the next page.<sup>11</sup> Now, as before, let nouns be driven into Num<sup>0</sup> by a strong feature in Romance. When YP is absent, then nouns will surface to the right of adjectives, both those in AP<sup>1</sup> and those in AP<sup>2</sup> 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

<sup>11</sup> Relevant to this decision is that ethnic adjectives can't appear prenominal 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  $\theta$ -role bearers are in nominals, this fact suggests that these external  $\theta$ -role bearers are positioned before objects in Romance nominals.



adjectives to have either interpretation. When YP is present, assume that  $Y^0$  has a strong feature, and the  $\text{Num}^0 + N^0$  pair will be driven into  $Y^0$ . In that situation, the noun will surface to the left of adjectives in  $\text{AP}^2$ , thus appearing to the left of adjectives with a restrictive interpretation, and will still remain to the right of adjectives in  $\text{AP}^2$ , presumably those with a non-restrictive interpretation.

Consider, by contrast, a language which does not allow Nouns to move into  $\text{Num}^0$  — Walloon, if Bernstein is right. In these languages, movement into  $Y^0$  will be blocked by the Head Movement Constraint. That is, if we could find something that not only prevented Nouns from surfacing in  $\text{Num}^0$ , but also prevented them from moving through  $\text{Num}^0$ , then we would have a way of correlating access to  $Y^0$  with access to  $\text{Num}^0$ .

This, in fact, is what Bernstein does. She argues that the plural morpheme in Walloon is placed in  $\text{Num}^0$ , and blocks movement of the noun in Walloon. That is, she claims that the plural morpheme resides in  $\text{Num}^0$  in Walloon and blocks not only liaison between the preceding adjectives and the following noun, but also blocks noun movement to  $Y^0$ . Let me briefly sketch the reasons she gives for this analysis.

The plural morphemes come in two forms: 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]. Illustrations are in (47).

## 5. Determiner Phrases and Noun Movement

---

- (47) a. les belè[s] feyes  
the pretty girls  
b. dès neûrè-z -amonnes  
some black berries

In singular nouns, neither of these morphemes appear:

- (48) 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, as in (49).

- (49) 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.

- (50) 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 I 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, as (51) below indicates.

- (51) a. Èle sont neûr.  
they are black  
b. Èle sont tot[es] petit[es].  
they are very little

And when the nominal that the adjective precedes is absent, this morpheme does not appear. It's missing in copular constructions, for instance, as shown in (52).

- (52) 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.

- (53) *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.

- (54) *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.

- (55) a. *grandeûr*  
 big  
 b. *grande amice [grāt amis]*  
 good friend

When adjectives are followed by the plural morpheme, they show obstruent final devoicing, as the contrast below illustrates.

- (56) a. \**grandè[s] fèyes*  
 big girls  
 b. *grantè[s] fèyes*  
 good girls

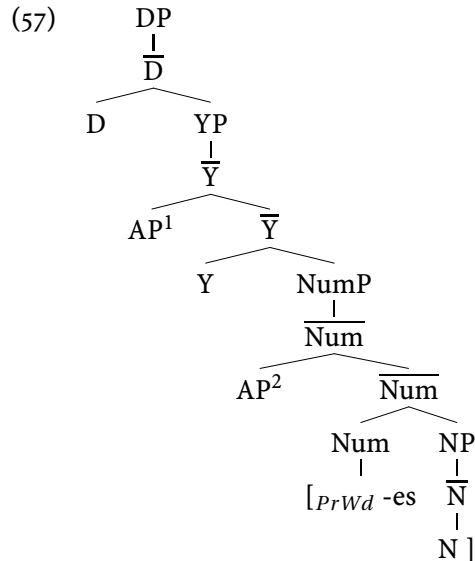
A second phonological reason for thinking that the plural affix is not part of the preceding adjective is that it is unstressed. She cites Morin who argues that all words in Walloon have stressed final syllables. Finally, again following Morin, she points to the fact that in Gondecourt Picard, the plural 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 next page.

## 5. Determiner Phrases and Noun Movement

---

(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

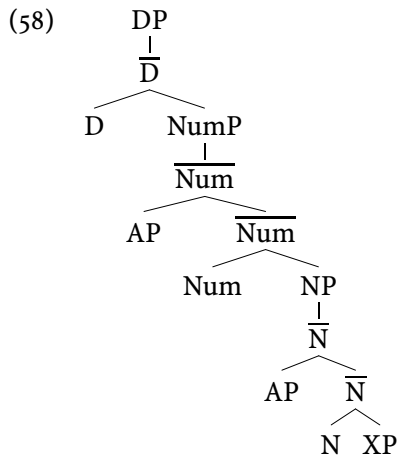


adjectives, because they will not be able to move through Num<sup>0</sup> into Y<sup>0</sup>.

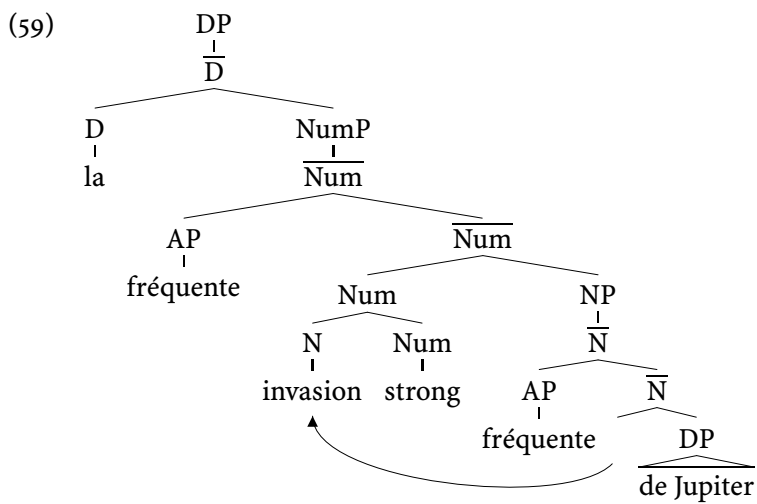
Still, there are problems with this account which are central enough to suggest that it needs revision. For one thing, it persists in requiring that adjectives be placed higher than NumP, and this runs against the trend Dryer found for languages to place adjectives within NumP. In addition, it credits the availability of a noun movement past an adjective to the properties of Y<sup>0</sup>, and only indirectly to the properties of Num<sup>0</sup>. But the aim of Bernstein’s analysis of the Walloon/French contrast is to make the properties of Num<sup>0</sup> responsible for noun movement past adjectives. Unless some intimate link can be made between Y<sup>0</sup> and Num<sup>0</sup>, the phrase marker in (47) isn’t equipped to express a correlation between occupying Num<sup>0</sup> and preceding single adjectives.

The decision to place adjectives higher than Num<sup>0</sup>, and to invent a new Y<sup>0</sup> into which nouns can move, responds to the desire to explain the absence of liaison in Walloon between prenominal adjectives and the nouns that follow them. Bernstein’s account forces Num<sup>0</sup> to intervene between prenominal adjectives and NP. Perhaps we should abandon trying to account for this fact, and let it come from some other idiosyncrasy of Walloon. This will allow us to return to a model of DPs like (58) below.





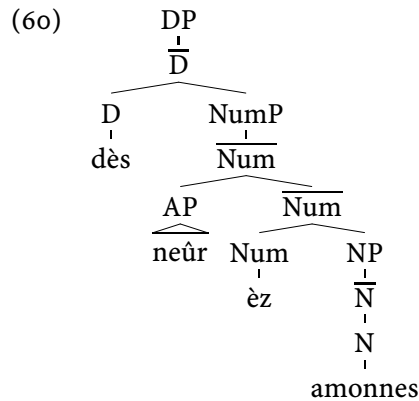
The difference between French and Walloon, as before, consists in whether Num<sup>0</sup> holds a free morpheme – a clitic – as in Walloon, or a number feature that matches the morphology on the following noun, as in French. Thus, in French we’ve got derivations like (59).



By contrast, in Walloon, Num<sup>0</sup> is occupied by a morpheme, not features, as in (60).

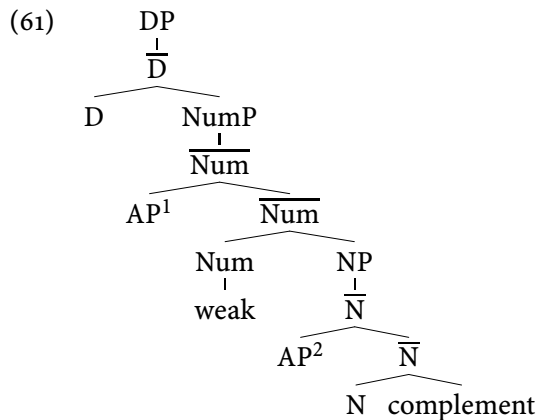
5. Determiner Phrases and Noun Movement

---

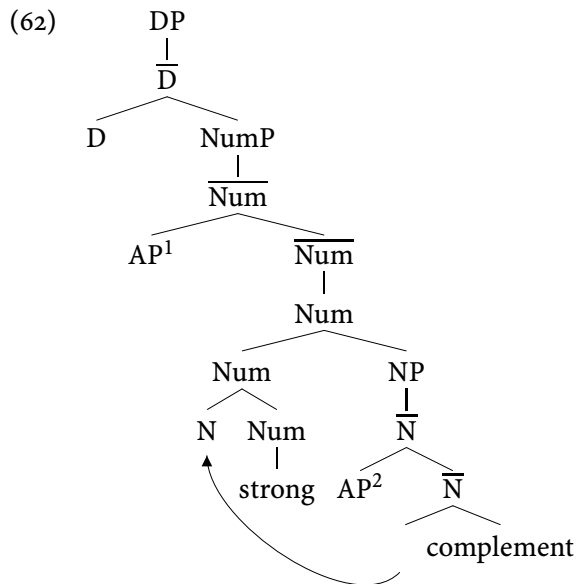


Perhaps the fact that the plural morpheme in Walloon must be immediately followed by a noun will block the existence of an adjective between  $\text{Num}^0$  and the noun. If this can be achieved, then Bernstein's suggestion that it's the presence of this plural morpheme between the adjective and the noun which blocks liaison between the adjective and noun can also be maintained. Perhaps this somewhat simpler model can explain all that Bernstein set out to account for, in other words.

Let's consider now this account of the difference in adjective placement between Walloon and French might be applied to the similar difference between English and French. There is no evidence of the sort we've just seen for Walloon that the number morpheme in English is a free morpheme. Let's assume, then, that  $\text{Num}^0$  in English contains features, as it does in French. Given the tools developed here, perhaps the most straightforward way of modeling the difference between English and Romance, then, would be to credit  $\text{Num}^0$  with strong features in French, Catalan and Spanish, but not in English. This will force nouns in French, Catalan and Spanish to move overtly to  $\text{Num}^0$ , thereby bringing them to the left of (some) adjectives, whereas in English this movement will necessarily be covert, given Earliest. So, English S-structures will fashion DPs as in (61) below.



In the remaining Romance languages, nouns will surface in Num<sup>0</sup> as in (62).



This gives us a three-way distinction. Walloon nouns don't move, English nouns move covertly and French nouns move overtly. The surface position of nouns in English and Walloon, then, is the same. But this arises for different reasons.

One consequence of forcing nouns in Walloon and English to remain in their underlying position is that they will remain to the right of the Specifier of NP. Recall that in Romance, we associated the ability of nouns to surface to the left of Specifier of NP with the availability of “subjects” of nouns to surface post-nominally. For instance, the French example in (15c), repeated below, arises by leaving *de chaque peintre étranger* ('of each foreign painter') in Specifier of NP and moving the noun, *portrait* ('portrait'), past it to the left.

- (15) le portrait de chaque peintre étranger de son enfant  
 the portrait of each painter foreign of his child  
 'the picture by each foreign painter of his child'

If nouns don't move to Num<sup>0</sup> in English or Walloon, we would expect these post-nominal subjects to be unavailable in both English and Walloon. We've already seen that this is the case for English. But, interestingly, it doesn't seem to be the case for Walloon.

Walloon does allow the N+Subject word order. Bernstein illustrates this with examples like (63).<sup>12</sup>

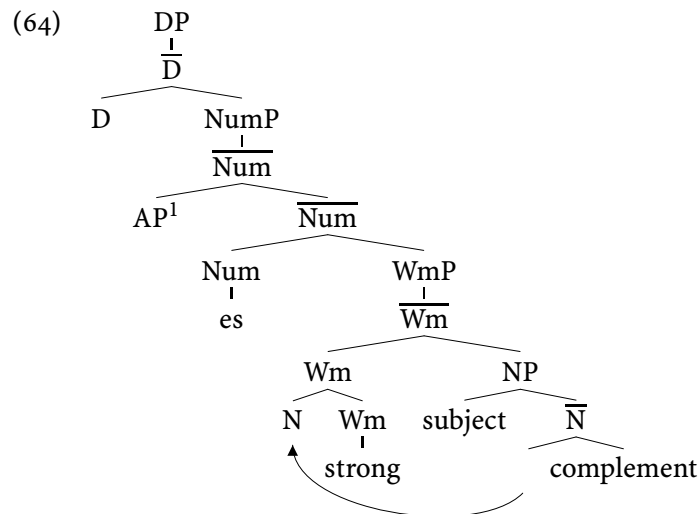
<sup>12</sup> This is perhaps not the most compelling example as it is difficult to tell whether *miller* bears the “subject” relation to *daughter*. Interestingly, Bernstein claims that Walloon also allows for postnominal

## 5. Determiner Phrases and Noun Movement

---

- (63) la fèy do mounî  
 the daughter of the miller  
 ‘the miller’s daughter’ (Bernstein 1993, (85): 241)

This suggests that even in Walloon, there is short noun movement, past the Specifier of NP position. If Bernstein’s arguments concerning how number morphology is expressed in Walloon is correct, this short noun movement can’t be to Num<sup>0</sup>. Bernstein suggests that it is instead movement to the position associated with the “gender” morpheme that Romance nouns so typically end in. She calls this a “word marker.” A schematic surface phrase marker for a Walloon DP, then, looks something like (64) below, then.



Note that this requires that adjectives are **not** capable of being within NP. Indeed, Walloon illustrates that the availability of postnominal subjects and postnominal (single) adjectives do not correlate cross-linguistically. It is necessary, therefore, to divorce the processes that yield these two word-orders, and if noun movement is the relevant process, then this means there must be two positions to which nouns can move, with subjects below, and adjectives above, the lower of these. Up to now we have relied on a general theory of modifier placement one of whose outcomes is that adjectives should be adjoined to the  $\bar{N}$  that they modify. One thing we learn from this study, then, is that this general theory will have to be changed.

But let’s leave that for a later occasion.

adjectives when they are ethnic; in fact, in these situations, the prenominal position is blocked. This also, rather dramatically, supports the conclusion that “subjects” can be postnominal in Walloon.

We must also revisit our account for why postnominal subjects don't arise in English. It's no longer sufficient to prevent nouns from moving to Num<sup>0</sup> in English. We must also now address the possibility that nouns could move to a position beneath Num<sup>0</sup>: the Wm<sup>0</sup> position that Bernstein posits Walloon to have, for instance.

While it might be that there are no postnominal subjects in English because English nouns don't make even a short move, it could also be because the other ingredient necessary to getting postnominal subjects is missing from English. Perhaps 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. Note that this Case is expressed in Romance with 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 "transitive" noun cannot be Case marked with *of*, as (65) indicates.

- (65) 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's not possible, note, even if these nouns do not express their object argument. Leaving the objects unexpressed in the examples in (65), for example, does not improve them.

- (66) a. \* the discussion of Jill  
 b. \* the of Jill discussion  
 c. \* the placement of Mark  
 d. \* the of Mark placement

But it's not 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 unaccusative verbs, as in (67).

## 5. Determiner Phrases and Noun Movement

---

- (67) a. the death of her  
b. the arrival of it  
c. the appearance of Jill  
d. the sinking of the ship

With nouns derived from unergative verbs, the situation is somewhat intermediate, as illustrated by (68) below.<sup>13</sup>

- (68) 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 (69).

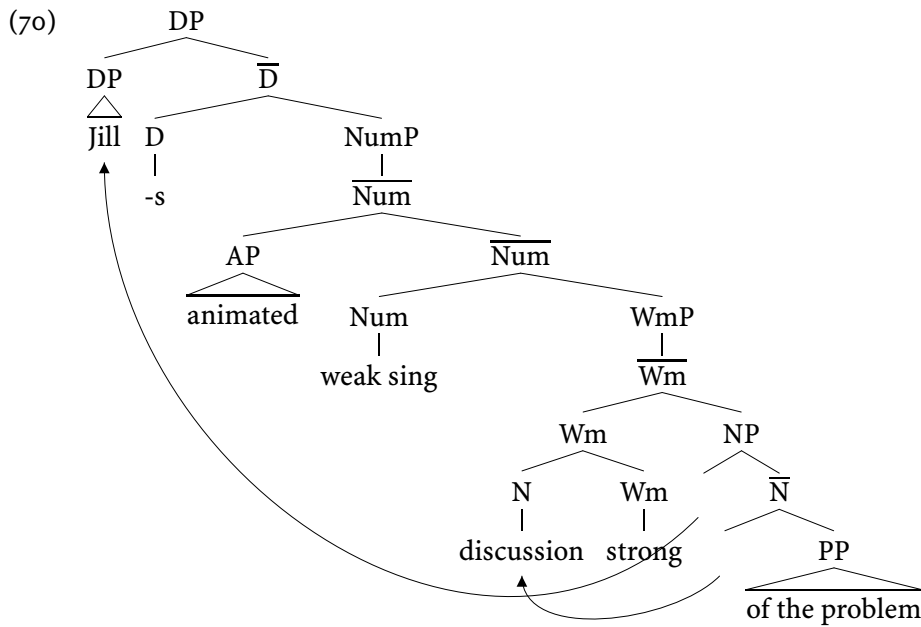
- (69) NOMINAL CASE ASSIGNMENT: ENGLISH  
Nominal Case is assigned to the “object position” of nouns.

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  $Wm^0$  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  $Wm^0$  it should be able to assign its Case to a DP within Specifier of NP: just the effect we are hoping to avoid.

It would be reasonable, therefore, to expect the general absence of postnominal subjects in English DPs to be caused by the constraints on Nominal Case that derive (69). This means it is conceivable that nouns in English do make a short movement, as they do in Walloon. This hypothesis, then, would give an example like “Jill's animated discussions of the problem” a representation like that in (70) on the facing page.

As this discussion makes clear, the relative heights of  $Wm^0$  and  $Num^0$  correlates the relative position of nouns and subjects with the relative position of nouns and adjectives. The general prediction is that there should be a positive correlation

<sup>13</sup> See Grimshaw (1990) for a discussion of these facts and an argument that nouns divide into the unaccusative and unergative classes.



between nouns surfacing to the left of (bare) adjectives and nouns surfacing to the left of subjects. We should not find<sup>14</sup> a language, in other words, that is the anti-Walloon: nouns surface to the left of bare adjectives but cannot surface to the left of “subjects.” In fact, this correlation does seem to hold in our language sample. All the languages we have examined that allow nouns to surface to the left of adjectives also allow them to surface to the left of subjects.

<sup>14</sup> Restricting attention to just those languages that Case mark subject DPs in Specifier of NP, and have the ban on right-adjoining bare adjectives.





# 6

## 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  $\theta$ -role assigner. In the case of clauses, the inflectional phrases are Agreement and Tense, and the  $\theta$ -role assigner is a verb and its projection. In the case of DPs, the inflectional phrases hold number morphology and something else, Bernstein speculated a "word marker," while the  $\theta$ -role assigner is a noun and its projection.

We came to this picture by examining, among other things, the position that "subjects" of DPs can take. In that discussion, I used the term "subject" to refer to the genitive marked DPs in examples such as (1).

- (1) a. Sean's opinion of poi
- b. Gary's novel of Mary
- c. Sandy's picture of Sean
- d. Mary's discussion of poi
- e. Mary's invasion of Jupiter
- f. Sandy's dancing
- g. Sean's running

In using this term "subject" in connection with clauses, I always referred to arguments of the predicate of that clause. Let's begin this chapter by considering whether this is also true for "subjects" of DPs. This will require us to investigate the argument structure of nouns which, as we will see, sheds light by way of comparison on the argument structure of verbs. This will eventually require a sweeping

## 6. Complement Structure

---

change to the picture we have developed for how verbs combine with their arguments.

### 6.1 Nouns and the $\theta$ -roles they assign

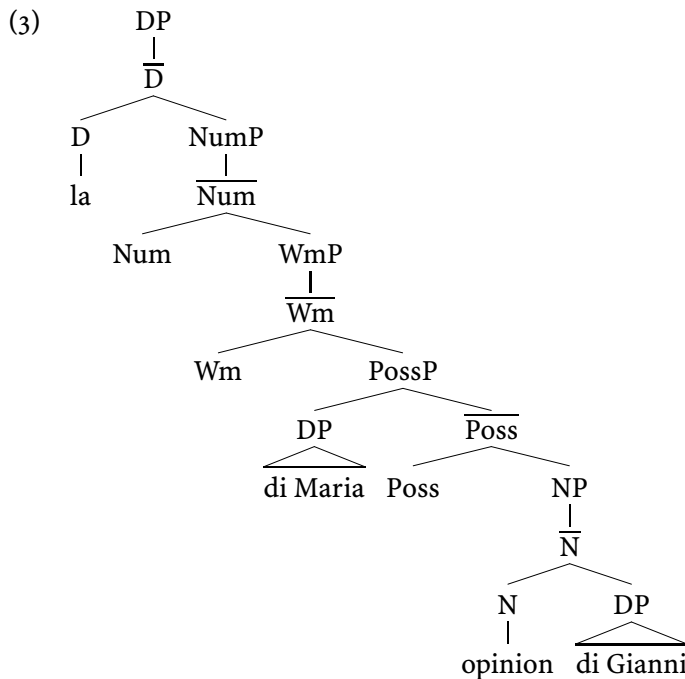
We saw that for one phrase to be an argument of another, its semantic role in the situation described by the sentence must depend on the denotation of the term it is an argument of. Our rough experimental technique for discovering arguments was to vary the term we suspect our candidate to be an argument of and see if its role changed as a consequence. By this criterion, I think we can see that the role played by the genitive DPs in (1a-c) is not determined by the meaning of the noun, or the  $\bar{N}$ , that follows. This is clearest, perhaps, for (1b) and (1c); in both cases, the relation the genitive bears to the rest of the DP is constant. We might call this relation “possession.” This very same relation is found in all cases in which the noun refers to an object:

- (2) a. Jerry’s cap, Smith’s finger, Mary’s doctorate, Sal’s friends, ...
- b. Jones’s regret, Smith’s desire, Mary’s opinion, Sal’s whim, ...
- c. Jerry’s intelligence, Smith’s size, Mary’s integrity, Sal’s fame, ...
- d. yesterday’s class, today’s mistakes, Monday’s headaches, ...

As can be seen from these examples, what “possession” means varies somewhat and does, indeed, vary with the head noun involved. But I think it can be said that in every case it is still a possession relation that is involved. Let us suppose that there is a hidden predicate in DPs that carries this meaning. Given that postnominal DPs in Romance can be possessors in this sense, we can speculate that this hidden predicate is lower than the surface position of moved nouns in Romance. A first approximation, then, might be a  $\mathcal{D}$ -structure parse like that in (3) on the next page for these sorts of cases. Understand *Poss* to be responsible for the possession relation that holds between the argument in its Specifier and the NP that follows. The surface form of this Italian example can then be generated by moving the noun into  $\text{Num}^0$  in the way described in the previous chapter.<sup>1</sup>

In the remaining examples of (1) the “subject” has a role that is identical to that which would be given to the external argument of the verb from which the head noun derives. In *Sandy’s dancing*, for instance, *Sandy* bears the same role that it does in *Sandy dances*; similarly, *Mary* bears the same role in *Mary’s discussion of*

<sup>1</sup> There is a fat literature addressing nature of *Poss*. Two of the central questions about it are: (1) does it have a single unit meaning, or are there are a variety of different relations that *Poss* names, and (2) is the meaning I’ve credited to *Poss* actually part of the denotation of the noun. Two good entry points to this literature are Barker (1995) and Partee and Borschev (2003).



*poi* and *Mary discussed poi*. In these cases it is necessary, then, to capture the fact that the same semantic relation is produced for both noun and verb, and using our present language this requires that we see the nouns in (1d-g) as assigning an external  $\theta$ -role to the “subject.”

Although less salient, it is also possible for the genitive DPs in (1d) and (1e) to have the possession relation. Consider the following scenario, for instance.

Bush, in an effort to boost his ratings, instructed the Pentagon to draw up plans for expanding the war on terror to its logical extreme. To Mary Sklar fell the unlucky job of figuring out the logistics of extending the war into space. Because of her considerable political and accounting skill, only Mary’s invasion of Jupiter was deemed too expensive.

In the last sentence of this paragraph, *Mary* is understood as the possessor of the invasion, and not the invader. This reading is aided by the fact that invasions rarely involve single individuals, and as a consequence we are prepared to jettison the more preferred Agent reading for the subject of this DP. When this aid is not present, it is more difficult to coax a possessive reading out; consider:

The Iron Chef’s A-team convened to discuss the best way of preparing and serving *poi*. This auspicious occasion was recorded by the repre-

## 6. Complement Structure

---

representative of the BBC: Mary Sklar. The week after, the Culinary Institute of America convened a similar conference, which was recorded by the representative of PBS: Sean Oggleby. ?Mary's discussion of poi was later broadcast on BBC 4.

With some effort, I think I can convince myself that *Mary* in the final sentence has the possession relation. But I'm not certain.

The examples in (1f) and (1g), by contrast, have only a reading for their "subject" in which it bears the  $\theta$ -role that would have been assigned by the verbs *run* and *dance*. This seems to be a generalization about nouns that are formed from verbs by *ing* suffixation, as we shall see.

In the situations where the "subject" of the noun bears the same  $\theta$ -role that would be expected from the related verb, let us imagine that this  $\theta$ -role is assigned to Specifier of NP by  $\bar{N}$ . This would make the conditions under which the external  $\theta$ -role is assigned the same for both nouns and verbs. As with the external  $\theta$ -role of verbs, the external  $\theta$ -role of nouns is determined by the head in combination with its complements. Thus we have shifts in meaning like those in (4), which mimic what we saw for verbs.

- (4) a. Jerry's throwing of the ball
- b. Jerry's throwing of the election
- c. Jerry's throwing of the party

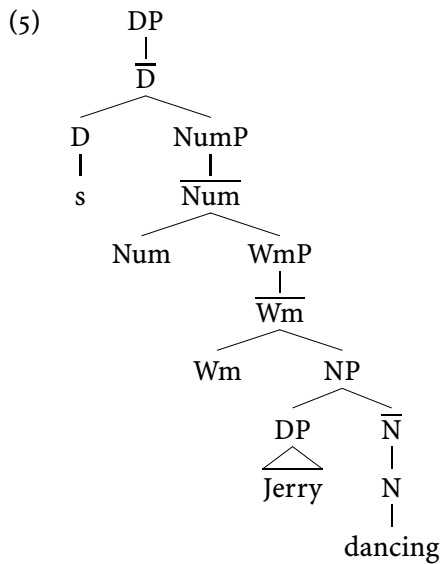
For these situations, then, we posit a D-structure like that in (5) on the facing page. The surface form of these English example is achieved by moving *Jerry* into the Specifier of DP, where it meets the Case filter's requirements. (The noun *dancing* might also move — into  $Wm^0$  — we haven't yet been able to determine this.)

If we concentrate on cases like those exemplified by (1d) and (1e), a question arises concerning the THETA CRITERION. The THETA CRITERION, recall, requires that terms which have  $\theta$ -roles assign those  $\theta$ -roles each to a unique "argument." If the noun *invasion*, then, assigns  $\theta$ -roles identical to the verb *invade*, why isn't the possessive reading for the genitive in (6a) as much a THETA CRITERION violation as (6b)?

- (6) a. Mary's invasion of Jupiter
- b. \* It invaded Jupiter.  
          where *it* is an expletive

It's the THETA CRITERION that requires *invade Jupiter* to assign a  $\theta$ -role, preventing an expletive from living in the subject position of (6b); why doesn't the THETA CRITERION similarly require *invasion of Jupiter* to assign a  $\theta$ -role in (6a)?

Similarly, why isn't (7a) ungrammatical for the same reason that (7b) is?



- (7) a. Mary's discussion  
 b. \* Mary discussed.

The THETA CRITERION forces the complement of *discuss* to be expressed in (7b), so why doesn't it have a similar consequence for (7b)?

Grimshaw (1990) argues that the solution to this problem in (7) relies on an ambiguity in the meaning of these nouns. Drawing on a distinction that goes back at least to Lees (1963), she notes that DPs with these nouns in them can refer to events of a certain sort,<sup>2</sup> or they can refer to individuals.<sup>3</sup> It is only on their eventive meaning that they name relations like verbs do, and consequently assign  $\theta$ -roles. On their other meaning they have the same status as the nouns in (1a-c). Grimshaw proposes that in cases such as (7a), the noun is not eventive, and does not therefore have  $\theta$ -roles to assign. It is only because the possession relation is so vague that it seems like the genitive in (7) appears to bear the external  $\theta$ -role we would expect the subject of the verb *discuss* to bear.

As support for this proposal, she endeavors to find ways in which to force the genitive in examples like (7a) to bear a  $\theta$ -role from the noun, and then determine

<sup>2</sup> She calls them "complex events," to distinguish them from the meaning that nouns like *dance* have; *the dance* can be said to refer to an event where dancing figured prominently.

<sup>3</sup> Lees called this distinction, or one roughly approximate to it, the difference between a noun referring to a "process" or the "result" of that process. These deverbal nominals are consequently sometimes called process or result nominalizations.

## 6. Complement Structure

---

if the result is grammatical depending on whether the object also appears. One of these techniques is to use modifiers that require events: *intermittent*, for instance. If we take a series of nouns that seem as though they can assign a subject  $\theta$ -role but no object  $\theta$ -role, as in (8), and add *intermittent* to them, the results are grammatical only if the object argument appears: witness (9).<sup>4</sup>

- (8) a. Mary's discussion  
b. Sal's toss  
c. Jerry's organization  
d. Smith's sale  
e. Jones's gift  
f. Sean's rental  
g. Jane's explanation
- (9) a. i. \* Mary's intermittent discussion  
ii. Mary's intermittent discussion of the syntax assignment  
b. i. \* Sal's intermittent toss  
ii. Sal's intermittent toss of the dishes  
c. i. \* Jerry's intermittent organization  
ii. Jerry's intermittent organization of the classroom  
d. i. \* Smith's intermittent sale  
ii. Smith's intermittent sale of limited edition *Sailors*  
e. i. \* Jones's intermittent gift  
ii. Jones's intermittent gift of gourmet poi  
f. i. \* Sean's intermittent rental  
ii. Sean's intermittent rental of scuba gear  
g. i. \* Jane's intermittent explanation  
ii. Jane's intermittent explanation of particle physics

Another kind of modifier with the same effect is a "rationale clause," which is an infinitival modifier that spells out the reasons some event took place; (10) illustrates.

- (10) Holly left to prove a point.

<sup>4</sup> It is important in this experiment to choose only nouns whose companion verbs require their objects. It should also be noted that in some of these cases – specifically (8b,d,e) – *intermittent* manages to be grammatical, even when no object is expressed, by modifying the time at which the nouns exist.

Rationale clauses should induce an eventive reading for nouns as well, then, and can be used to test whether the THETA CRITERION holds when nouns have this interpretation. As Grimshaw's thesis predicts, adding a rational clause to the DPs in (8) leads to ungrammaticality unless the object is also expressed.

- (11) a. i. \* Mary's discussion to impress her roommate  
 ii. Mary's discussion of the syntax assignment to impress her roommate  
 b. i. \* Sal's toss  
 ii. Sal's toss of the dishes to scare away his guests  
 c. i. \* Jerry's organization to instigate a riot  
 ii. Jerry's organization of the classroom to instigate a riot  
 d. i. \* Smith's sale to bring in some needed cash  
 ii. Smith's sale of limited edition *Sailors* to bring in some needed cash  
 e. i. \* Jones's gift to increase the variety of dishes  
 ii. Jones's gift of gourmet poi to increase the variety of dishes  
 f. i. \* Sean's rental to add income  
 ii. Sean's rental of scuba gear to add income  
 g. i. \* Jane's explanation to entertain us  
 ii. Jane's explanation of particle physics to entertain us

Finally, as noted above, the *ing* forms of deverbal nouns resist having a possession reading for their "subjects." And in these cases too, the object is obligatory.

- (12) a. i. \* Mary's discussing  
 ii. Mary's discussing of the syntax assignment  
 b. i. \* Sal's tossing  
 ii. Sal's tossing of the dishes  
 c. i. \* Jerry's organizing  
 ii. Jerry's organizing of the classroom  
 d. i. \* Smith's selling  
 ii. Smith's selling of limited edition *Sailors*  
 e. i. \* Jones's giving  
 ii. Jones's giving of gourmet poi  
 f. i. \* Sean's renting  
 ii. Sean's renting of scuba gear

## 6. Complement Structure

---

- g. i. \* Jane's explaining
- ii. Jane's explaining of particle physics

Let's tentatively adopt this way of accounting for the apparent THETA CRITERION violation in cases where the object is missing, as in (7). When nouns have  $\theta$ -roles to assign, they obey the THETA CRITERION; it's only because nouns are systematically able to have a meaning in which they don't assign  $\theta$ -roles that leads to the contrary appearance.

Let's turn next to cases where the THETA CRITERION appears to be violated because of a missing subject argument, as in (6). Other, simpler, examples of this problem can be seen now in examples like (13).

- (13) a. the discussing of the syntax assignment
- b. the tossing of the dishes
- c. the organizing of the classroom
- d. the selling of limited edition *Sailors*
- e. the giving of gourmet poi
- f. the renting of scuba gear
- g. the explaining of particle physics

One popular solution to this case resorts to the same device that the parallel problem in infinitival clauses uses: PRO.<sup>5</sup> In the case of infinitival clauses, we arrived at the conclusion that they contain a silent subject argument, i.e. PRO, by way of the following reasoning. First, we noted that these infinitival clauses are ungrammatical unless an external  $\theta$ -role is assigned in them. This can be explained by way of the THETA CRITERION if these infinitival clauses are forced to have PRO in them. The EPP forces something to occupy Specifier of AgrP, and so in infinitival clauses, the EPP will have the effect of forcing PRO in them. Thus, the EPP leads to the expectation that there will be an argument in infinitival clauses, contrary to what is visible. Being good linguists, we trust our principles over our perceptual apparatus.

That argument cannot be transferred to DPs, however, because the EPP's effects are not felt here. There is no requirement, in other words, that forces the Specifier of DP (or any other Specifier within DP) to be filled. DPs do not host expletives.

There are other ways of seeing PRO, however; most of these we will not encounter until towards the end of these lectures.

To foreshadow one of these that can be used in this context, consider the reflexive pronoun *oneself*. Like other reflexive pronouns, *oneself* requires that there be something else in its sentence that it corefers with. This other term is known as its

<sup>5</sup> Roeper (1985, 1987) is the *locus classicus* for this position.



“antecedent.” It’s this requirement that is responsible for the contrast in grammaticality in (14).

- (14) a. Jerry only talks to one about oneself.  
 b. \* Jerry only talks about oneself.  
 compare: *Jerry only talks about poi.*

There is no antecedent for *oneself* in (14b) and this causes it to be ungrammatical. Examples like (15) indicate that PRO can be an antecedent for *oneself*.

- (15) a. To talk about oneself is appropriate in a therapy session.  
 b. To incessantly groom oneself amuses Jerry.

Putting these two observations together with the grammaticality of (16) permits an argument for the presence of PRO in the subject position of DPs.

- (16) a. The incessant discussing of oneself is not recommended.  
 b. The tossing of oneself against the door is not going to help.  
 c. The selling of oneself is generally frowned upon.  
 d. The giving of oneself to a good cause is not encouraged by capitalist economies.  
 e. The renting of oneself is not the same as selling one’s labor.

One solution, then, is to conclude that DPs whose nouns do assign  $\theta$ -roles, and have no overt subject argument, obligatorily have PRO as their subject. To achieve this outcome, we will have to design our theory that controls the distribution of PRO to allow PRO to stand within DPs. We shall encounter soon a rethinking of the distribution of PRO that makes this difficult.

If PRO is not present in these DPs, then we must find another explanation for the apparent violation of the THETA CRITERION with respect to subject arguments. Another direction that Angelika Kratzer has suggested is that the way in which subject and object  $\theta$ -roles are assigned differs in a way that allows them to come apart in nominals. The approach would allow nouns to assign object  $\theta$ -roles without assigning subject ones. On this view too, then, there is no violation of the THETA CRITERION in these examples. We’ll return to this alternative approach.

There is considerable work to be done, of course. But let us continue as though nouns can, at least on one of their meanings, assign  $\theta$ -roles under the same conditions that verbs do. And, as many of the examples we have looked at illustrate, it is generally true of deverbal nouns that they inherit their argument structure from the verbs they are derived from. What we should expect from these observations, then, is that the ways in which arguments of verbs are expressed should match the ways in which the arguments of eventive nouns are expressed.

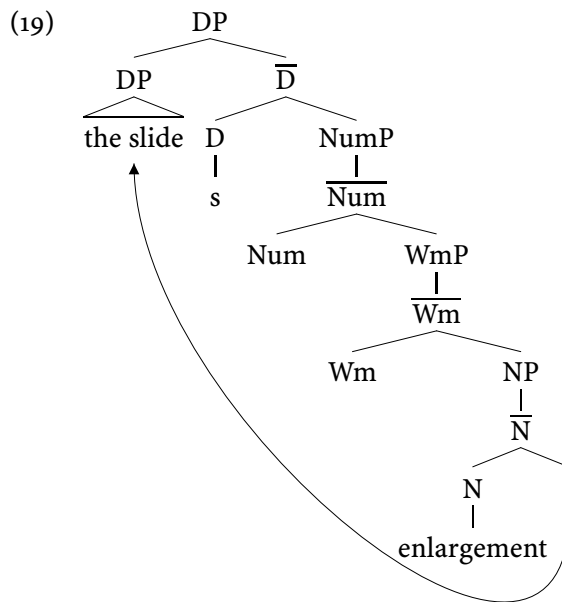
## 6. Complement Structure

---

One of the things we discovered about how arguments of verbs are expressed is that they sometimes undergo Argument Movement. There is evidence that this also happens to the arguments of nouns. We have seen, for example, that there is an operation, PASSIVE, which sets up situation in which the object of a verb, if it should need to satisfy the Case filter, moves into the Nominative case marked Specifier of AgrP. There are DPs that have a very similar shape to them. For example (18) bears a superficial similarity to what the PASSIVE would form from (17).

- (17) a. The Iberians' construction of Rome.  
 b. Barry's enlargement of the slide.
- (18) a. Rome's construction (by the Iberians)  
 b. the slide's enlargement (by Barry)

Preserving the match between the syntax of verbal arguments and nominal arguments, these nominals would therefore have a representation something like that in (19) below, once the "object" has moved.



There are ways in which the argument structure of nominals is not like that of clauses, however. Many of these are explored in a famous paper by John Ross called "Nouniness."<sup>6</sup> One of these differences concerns the typology of infinitival

<sup>6</sup> See Ross (1974).

clauses that are found as complements to verbs. Recall that these infinitives partition into two classes: permeable and impermeable. The impermeable ones can house PRO, prevent A chains from spanning them and don't allow Case assignment into their "subject position." Permeable infinitives have the opposite mix of properties: they cannot house PRO, they 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: impermeable infinitives are CPs and permeable ones are AgrPs.

Interestingly, nouns do not seem to permit permeable infinitives. This can be appreciated by noting that verbs which select impermeable infinitives continue to do so when they are nominalized, as in (20).

- (20) a. i. I attempted [to read all of *Aspects*].  
       ii. my attempt [to read all of *Aspects*].  
       b. i. I promised [to start reading *LSLT*].  
       ii. my promise [to start reading *LSLT*].  
       c. i. I persuaded Sal [to read *Connectedness*].  
       ii. my persuasion of Sal [to read *Connectedness*].

But, so far as I know, no verb which selects a permeable 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 (21).

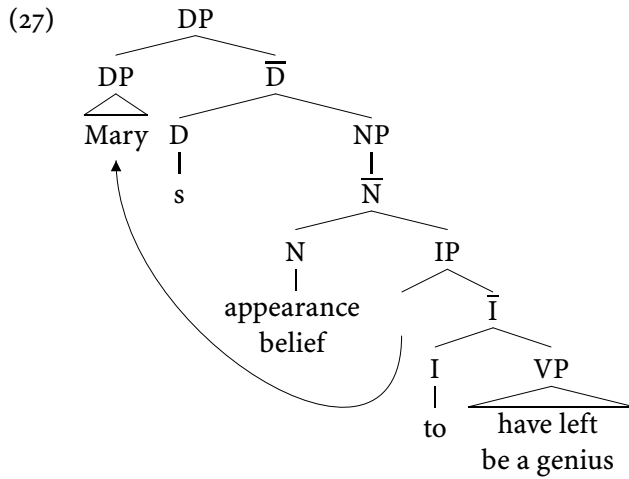
- (21) a. i. I believe [Sal to be happy].  
       ii. \* my belief [of Sal to be happy].  
       b. i. I consider [Sandy to be intelligent].  
       ii. \* my consideration [of Sandy to be intelligent].  
       c. i. I find [Jeri to be a good actress].  
       ii. \* my finding [of Jeri to be a good actress].

And verbs whose complement infinitives have an argument that moves out of that infinitive, are also ungrammatical when nominalized.

- (22) a. i. I appear [to be confused].  
       ii. \* my appearance [to be confused].  
       b. i. This proposal is likely [to be incomplete].  
       ii. \* this proposal's likelihood [to be incomplete].  
       c. i. She seems [to be tired].  
       ii. \* her semblance [to be tired].

The absence of this last class of permeable infinitive in nominals can be seen too by observing that the "passive" guise of the nominals in (21) are also ungrammatical.





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  $\theta$ -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.<sup>8</sup>

This leaves us still with the task of explaining why the other form of permeable 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 island-hood of complements to nouns.

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 (28) illustrate this generalization.

<sup>8</sup> That is, these phenomena support treating the long-distance dependency that (25) illustrates arise in questions in English with the same mechanism that is used to treat "Argument spread."

## 6. Complement Structure

---

- (28) a. i. I believe this fact interesting.  
ii. \* this fact's belief interesting.  
b. i. I consider this sentence instructive.  
ii. \* this sentence's consideration instructive.  
c. i. Some find these facts remarkable.  
ii. \* 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 (29).

- (29) 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 (28).

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 (28) 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 (30) are also ungrammatical.

- (30) 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 permeable infinitives, and when we find the reason Case assignment by nouns is blocked into permeable infinitives, we can spread that account to the cases in (30).

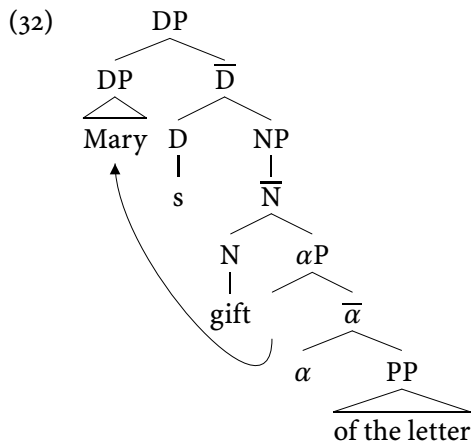
Another place where nouns don't seem capable of supporting the complement structure that verbs can is found with the double object construction. We have put off understanding how these constructions can be squared against Stowell's hypothesis that what forces DPs to be initial among complements is their need to be adjacent to the Case assigning verb. These constructions, then, are presently just puzzles for our theory. The fact that double objects are not possible as complements to nouns, as (31) illustrates, perhaps offers information that might help us understand how to treat them.

- (31) a. i. Her teacher gave Mary the letter.  
ii. \* Her teacher's gift of Mary of the letter.  
iii. \* Mary's gift of the letter (by her teacher).

- b. i. Her classmate offered her a crayon.
- ii. \* her offer of a crayon by her classmate.
- iii. \* her classmate's offer of her of a crayon.
- c. i. This salesman sold us that defective natto pot.
- ii. \* our sale of that defective natto pot (by this salesman).
- iii. \* this salesman's sale of us of that defective natto pot.
- d. i. The coach tossed him the ball.
- ii. \* his toss of the ball (by the coach).
- iii. \* the coaches toss of him of the ball.
- e. i. I rented John the office.
- ii. \* John's rental of the office (by me).
- iii. \* my rental of John of the office.

In a pattern that is now familiar, the double object construction appears to be a sort of permeable 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 (32).



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 (33).

## 6. Complement Structure

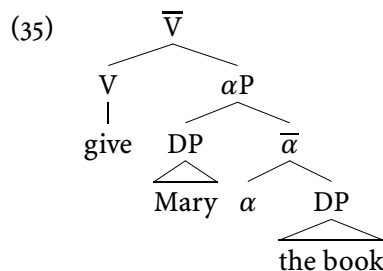
---

- (33) 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:

- (34) 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 (35).



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 (1984c). Extending the treatment of the absence of permeable 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.



## 6.2 Double Object constructions and Larsonian shells

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, and 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:

(36)  $\alpha$  is in the scope of  $\beta$  iff  $\beta$  c-commands  $\alpha$ .

The examples in (37) illustrate a variety of scope sensitive phenomena.<sup>9</sup>

- (37) a. Mary showed the boys<sub>1</sub> each other's<sub>1</sub> parents.  
 b. \* Mary showed each other's<sub>1</sub> parents the boys<sub>1</sub>.  
 c. Mary gave each boy<sub>1</sub> his<sub>1</sub> toy.  
 d. \* Mary gave its<sub>1</sub> owner each toy<sub>1</sub>.  
 e. \* Mary baked her<sub>1</sub> Sally's<sub>1</sub> recipe.  
 f. Mary baked Rover's<sub>1</sub> owner Rover<sub>1</sub>/the dog<sub>1</sub>.

In (37a) and (37c), the expression *each other* and *his* are referentially dependent on *the boys* and *each boy* in a particularly direct fashion. The semantic value that these terms gets is determined by the semantic values assigned to these preceding DPs. In (37a), for instance, the individuals that *each other* refers to is drawn from the same set of individuals that *the boys* does. Suppose that there are just three boys: Tom, Max and Sean. Then (37a) has a meaning that can be equivalently expressed with (38).

- (38) Mary showed Tom Max and Sean's parents, and Mary showed Max Tom and Sean's parents, and Mary showed Sean Max and Tom's parents.

The referent of *each other* varies, then, in a way that depends on the referent of its “antecedent”: *the boys*.

Something similar is found in (37c), whose meaning, in our temporarily boy-impooverished world can be equivalently expressed with (39).

- (39) Mary showed Tom Tom's parent, and Mary showed Max Max's parents, and Mary showed Sean Sean's parents.

<sup>9</sup> This catalogue of scope asymmetries is explored in Barss and Lasnik (1986).

## 6. Complement Structure

---

Again, the semantic value that *his* receives varies as a function of the semantic value given to its antecedent: *each boy*. When a pronoun gets this reading, it is said to be a variable “bound” to its antecedent. Bound variable pronouns and anaphors must be within the scope of their antecedents.

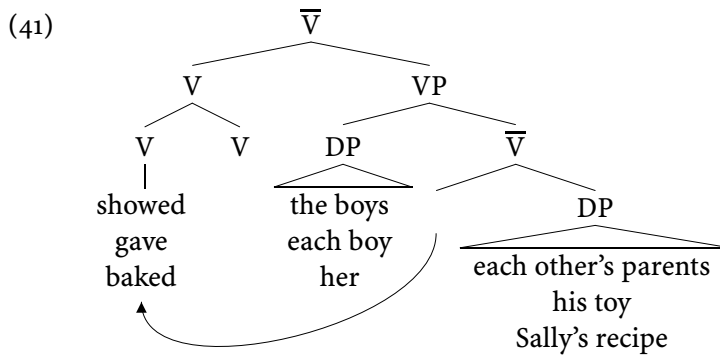
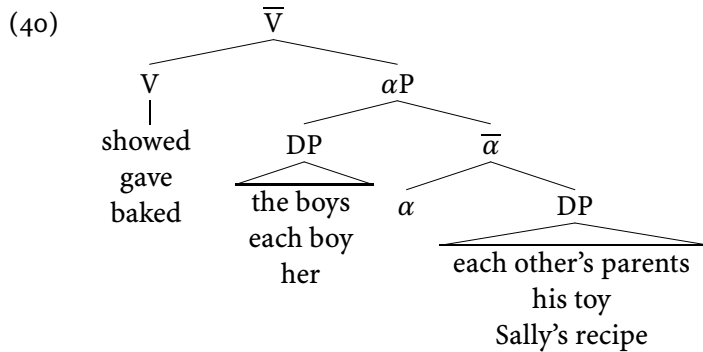
The referential dependence that *each other* and *his* have on their antecedents will be graphically indicated by way of numerical subscripts. The fact that *each other* and *the boys* both bear the subscript “1” is meant to indicate that *each other* is referentially dependent in the way described on *the boys*. Similarly, that *each boy* and *his* both bear the subscript “1” in (37c) is meant to indicate that *his* gets an interpretation in which it is a variable bound to *each boy*. There are some treatments of these semantic facts that credit syntactic representations with subscripts of these sorts and provide a way of interpreting the subscripts so that the intended interpretations emerge. At present, however, it is enough to use these subscripts merely as a typographical way of indicating what kind of meaning is intended.

In (37e) we see a kind of enforced referential independence between the terms *her* and *Sally*. If *her* is taken to be referentially dependent on *Sally* – indicated by coindexing them – the result is ungrammatical. This sentence can only be interpreted in such a way that *her* and *Sally* refer to independent individuals. We say in such cases that a “disjoint reference” effect arises.<sup>10</sup> This disjoint reference effect arises between a name-like DP and any other DP it is in the scope of.

What the contrasts in (37) indicate, then, is that in the double object construction, the second DP falls within the scope of the first DP, but the first DP does not fall within the scope of the second. A pronoun or reciprocal in the second DP can be referentially dependent on the first DP, but the second DP cannot serve as antecedent for pronouns or reciprocals in the first DP. And, similarly, a name-like term in the second DP must be disjoint in reference from the first DP, but a name-like term in the first DP need not be disjoint from the second. If scope is computed on c-command relations in the way that (36) describes, then what is required is giving the double object construction a representation like (40) on the next page. Of course, this is the same representation that we converged on in the previous section through a consideration of the facts of complementation in DPs.

Larson (1988) defends the parse for double objects in (40) and provides, in addition, a suggestion as to the identity of “ $\alpha$ .” He suggests that  $\alpha$  is, in fact, the verb, and that the higher VP — a VP he calls a “VP Shell” — is headed by an empty  $V^0$  position into which the verb moves. He suggests then that (40) has the representation in (41).

<sup>10</sup> See Lasnik (1976), where this way of thinking about such cases is first discussed.



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.<sup>11</sup>

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 "productive" in English, but he suggests that this is a special property of English and makes some suggestions as to what this property is.

<sup>11</sup> Larson's paper predates the Pollockian revolution to clausal structure, and so he identifies  $T^0$  with  $I^0$ .

## 6. Complement Structure

---

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 Mark Baker's Uniformity of Theta Role Assignment Hypothesis (UTAH), which states:<sup>12</sup>

- (42) UNIFORMITY OF  $\theta$ -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  $\theta$ -roles involved in both frames are 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 an operation in Categorical Grammar that Emmon Bach dubbed "Right Wrap."<sup>13</sup> This hypothesis 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  $\theta$ -role borne by the direct object, as in (43).

- (43) 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  $\theta$ -role assigned to the subject. We concluded in the case of the subject that the  $\theta$ -role assigned came from a phrase made up of the verb and its complement. In this situation, then, we might want to reach the same conclusion: the verb and indirect object form a phrase that together  $\theta$ -marks the direct object. This suggests, then, that there is a phrase that holds the verb and indirect object but excludes the direct object.

<sup>12</sup> See Baker (1988) for this proposal; UTAH is Baker's name for it. UTAH is one way of formulating an idea, sometimes called the "universal base hypothesis" that goes back several decades in Generative syntax; see Katz and Postal (1964). A closely allied principle, and one that probably influenced Baker's work, is the Universal Alignment Hypothesis of Perlmutter and Postal (1984).

<sup>13</sup> See Bach (1984).

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 which are discontinuous; one such case is (44).

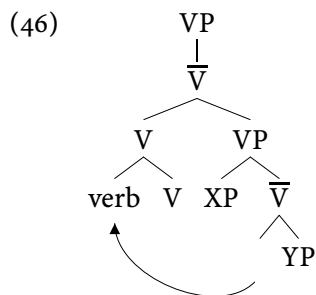
- (44) Some gave books to Sam, and others, magazines.

To the extent that syntactic operations like Gapping are constrained to apply only to constituents, we should conclude from cases such as these, 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.

Further, as Jackendoff reminds in a rejoinder to Larson’s paper,<sup>14</sup> there are scope asymmetries of the sort we’ve seen in double object constructions for all double complement verbs.

- (45) a. Mary showed every paper<sub>1</sub> to its<sub>1</sub> author.  
 b. \* Mary showed her<sub>1</sub> problem to every author<sub>1</sub>.  
 c. Mary introduced the girls<sub>1</sub> to each other’s<sub>1</sub> parents.  
 d. \* Mary introduced each other’s<sub>1</sub> parents to the girls<sub>1</sub>.  
 e. Mary told the fink<sub>1</sub> about her<sub>1</sub> shortcomings.  
 f. \* Mary told her<sub>1</sub> about the fink’s<sub>1</sub> shortcomings.

For these reasons, then (and one more we come to immediately), Larson suggests that the structure for ALL double complement verbs is as indicated in (46).



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

<sup>14</sup> See Jackendoff (1990).

## 6. Complement Structure

---

reconcile Larson's decision to treat all two complement constructions as involving an embedded small clause with these data, then. Let's come back to this.

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  $\theta$ -roles. Indeed, Larson's idea is that  $\theta$ -roles arrange themselves in a hierarchy that governs the position in the tree that they will be assigned to. The higher the  $\theta$ -role on the hierarchy, the higher in the phrase marker that  $\theta$ -role will be assigned. In particular, he suggests that AGENT is higher than THEME which is higher than GOAL/BENEFACTIVE which is higher than LOCATIVE, and so on.<sup>15</sup>

The final reason Larson offers for accepting this view of the arrangement of complements is that it provides a method for characterizing Ross's HEAVY NP SHIFT that helps explain some of its puzzling properties. One puzzle is why leftward movement, but not HEAVY NP SHIFT, can strand prepositions, as the contrast in (47) indicates.

- (47) 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 HEAVY NP SHIFT is not:

- (48) a. Sally, I gave the pictures.  
b. \* I gave the pictures all my relatives in Duluth.


And, finally, it is widely thought that HEAVY NP SHIFT is subject to constraints which govern the distance the moved term can go that are not found in other movement transformations. The other phrasal movement operation we have investigated is Argument Movement, and it has the ability to move arguments quite far. If the other conditions on Argument Movement are obeyed, it can take an argument out of several clauses, as illustrated by examples like (49), for instance.

- (49) Mary appears to seem to be believed to be unhappy.  
↑ \_\_\_\_\_

Other movement operations that relocate phrases have this ability as well. For instance, the process that moves interrogative phrases to form the questions in (25) and (26) — so-called *wh*-Movement — can take phrases considerable distances. In

<sup>15</sup> What's wanting, then, is an explanation for why this particular arrangement of  $\theta$ -roles should map onto the categories that it seems to.

(50), for instance, *which woman* has moved from a position next to the  $\bar{A}$  *unhappy* where its  $\theta$ -role derives.

(50) Which woman does it appear that you believe seemed to be unhappy?  


John Ross, when he introduced HEAVY NP SHIFT, argued by way of contrasts like (51) that it was not capable of relocating phrases out of the clause they originate in.

- (51) a. Sally decided [that Sam should give to her all the gourmet poi] on Tuesday.  
 b. ?? Sally decided [that Sam should give to her] on Tuesday all the gourmet poi.

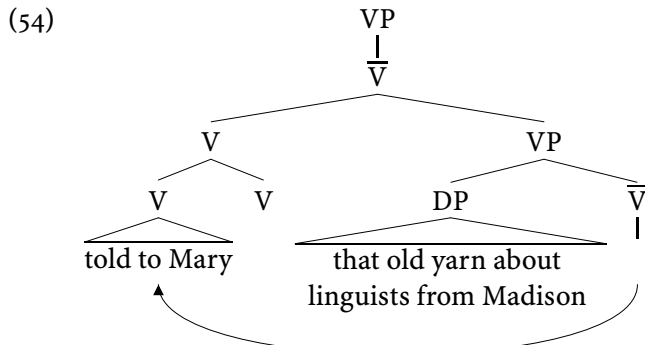
It is very difficult to understand (51b) in such a way that *on Tuesday* modifies the root clause; the overwhelming temptation is to understand it so that *on Tuesday* modifies the *give*-clause instead. This would follow if HEAVY NP SHIFT cannot move anything out of the clause it originates in.<sup>16</sup>

Larson's proposal allows for a reanalysis of HEAVY NP SHIFT that explains its differences from leftward movement. Indeed, instead of seeing HEAVY 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:

- (52) When  $\bar{V}$  contains one undischarged internal  $\theta$ -role, it may be reanalyzed as a  $V^0$ .

On this view, then, a HEAVY NP SHIFT example like (53) comes about in the way sketched in (54).

(53) Jerry told to Mary that old yarn about linguists from Madison.



<sup>16</sup> Ross argued that the constraint was even more draconian, preventing HEAVY NP SHIFT from moving anything out of the smallest maximal projection it originates in. We have already encountered, however, evidence from German that this is not cross-linguistically the case.

## 6. Complement Structure

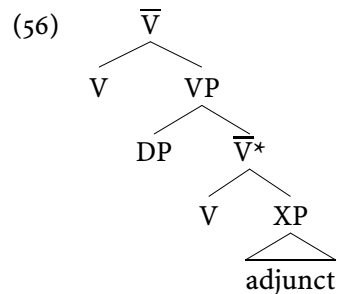
---

This straightforwardly explains the inability of HEAVY 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, HEAVY NP SHIFT does not strand prepositions because it never targets the complement of a preposition. Further, the apparently tighter conditions on how far HEAVY NP SHIFT can move things is derived as well. Because the “Heavy NP Shifted” item is not moving, but the predicate that that item combines with is, the distance that it appears the Heavy NP Shifted item moves will amount to the size of this predicate and how far it moves. The predicate moves as a head, on Larson’s view, and so it will be subject to the constraints on Head Movement. In English, this means the predicate will not be moving very far at all. Examples like Ross’s (51) are blocked because the string that would have had to have moved to the left of the object doesn’t even form a constituent. Finally, how this proposal derives the fact that the first object of a double object construction cannot HEAVY NP SHIFT will have to wait for us to see the structure he assigns to double object constructions.

If Larson’s characterization of HEAVY NP SHIFT is correct, then cases like the following suggest that we’ve got the position of adverbs screwed up.

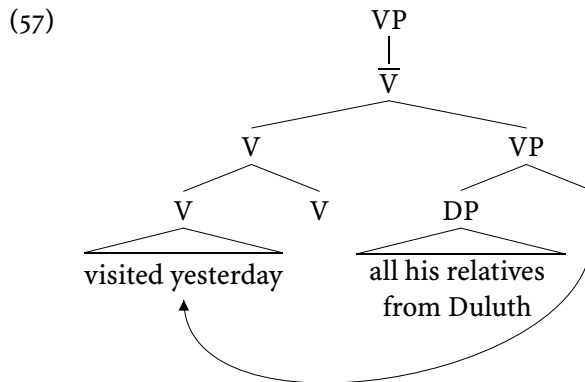
(55) Sam visited yesterday all his relatives in Duluth.

On Larson’s characterization of HEAVY 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  $\theta$  Hierarchy. So he gives a parse like that in (56) for cases where a verb combines with a DP complement and a non-argument.



His reanalysis rule will be able to affect  $\bar{V}^*$ , and thereby move verb and adjunct past the heavy DP in Specifier of VP. So the derivation behind (55), for instance, will be as in (57).





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  $\theta$ -role bearers closer to their  $\theta$ -role assigners than non  $\theta$ -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 HEAVY 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. He produces examples in which the first argument seems to have a non-argument in its scope, in much the same way that we’ve seen the first argument to have the second in its scope.

- (58) a. John visited few friends any day this week.  
(Larson 1990, (52):621)
- b. \* John filed them<sub>1</sub> without reading [Mary’s articles]<sub>1</sub>  
(from Contreras (1984), in Larson 1990 (53a):622)
- c. I will visit every child<sub>1</sub> before his<sub>1</sub> mother.

In (58a), the so-called “negative polarity item” *any* falls within the scope of *few friends*. We can deduce this from the fact that quite generally negative polarity items must be found in the scope of “negative” operators – like *few NP* – for them to be grammatical. This is the reason for the contrast in (59), for example.

- (59) a. Few people eat anything.  
 b. \* Anyone eats few foods.

## 6. Complement Structure

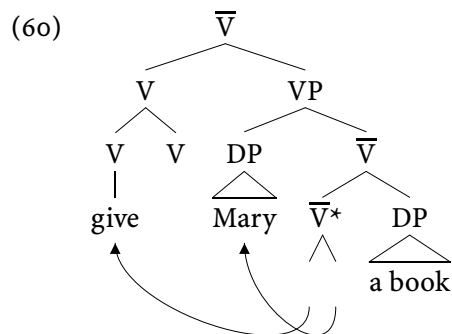
Only in (59a) is the negative polarity item c-commanded by *few foods*. In (58b) we have an instance of disjoint reference holding between *Mary's articles* and *them*. And in (58c) we have an example of the pronoun *his* being bound by *every child*.

In all these examples, then, 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 HEAVY NP SHIFT operation and scope facts 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 HEAVY 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 HEAVY 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 Specifier of VP position. (Note, then, that this operation must rob the verb only of what Larson would call its inherent Case.) Thus, this operation forms (60) from the DP+PP frame.



How does the demoted DP – *a book* – get Case? Larson suggests that 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, will allow *a book* to be governed by a Case assigning verb. Thus, the argument in Specifier of VP gets structural accusative Case, while the second argument DP

gets an inherent Accusative 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 only 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.

Any account that relates the DP+PP frame to the double object frame by rule is going to have to explain why some verbs allow one of these but not the other. For instance, *donate*, whose meaning is rather close to *give* has only the DP+PP frame, as (61) indicates.

- (61) a. She donated her cat to the charity.  
 b. \* She donated the charity her cat.

And *spare*, by contrast, allows only the double object frame.

- (62) a. \* She spared the ordeal to Fluffy.  
 b. She spared Fluffy the ordeal.

Larson charges the part of his rule that robs Case from the indirect object with this job. Suppose, he suggests, that *to* assigns a very particular GOAL  $\theta$ -role: “Goal of motion along some Path.” Further, imagine that it can be deleted by this process only when the  $\theta$ -role assigned by the verb is identical. Thus, he supposes, cases which allow only the DP+PP frame, such as *donate* or the other examples in (63), are those where the verb does not assign an identical  $\theta$ -role.

- (63) a. i. I distributed apples to the children.  
 ii. \* I distributed the children apples.  
 b. i. I contributed my time to the auction.  
 ii. \* I contributed the auction my time

Cases where only the DP+DP frame is possible, as with *spare* or *deny*, he suggests arise because the  $\theta$ -role that *to* would assign is not appropriate.

- (64) a. They denied him tenure.  
 b. \* They denied tenure to him.

Clearly if *to* can only assign a “Goal of motion”  $\theta$ -role, then this is not consonant with the meaning of these sentences.

The factors that control this alternation probably range beyond the role that *to* (or *for*) play in the DP+PP frame. In an exhaustive study of this alternation, Pinker (1989) and Gropen, Pinker, Hollander, Godberg, and Wilson (1989) found that a range of factors played a role, including whether the verb derives historically from the Germanic or Latinate parts of the English vocabulary. The difference between

## 6. Complement Structure

---

*give* (Germanic) and *donate* (Latinate) falls in line with the generalization that the Latinate verbs do not participate in this alternation.<sup>17</sup> It is difficult to see the difference in meaning between these two verbs that Larson's proposal would require.

In any case, as far as I can see, Larson's proposal 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  $\theta$ -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  $\theta$ -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).

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 HEAVY 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 internal  $\theta$ -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 (60). 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 internal  $\theta$ -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  $\theta$ -role from the trace in sister of  $V^0$  position, and the THEME  $\theta$ -role is borne by the DP adjoined to  $\bar{V}$ . So the restriction in (52) would prevent reanalysis of  $\bar{V}^*$ .

Let's summarize what we've seen in Kayne's and Larson's work. I'll list below the conclusions Larson has reached and for each list the reasons that he gives, or that we've seen in the nominalization data, for these conclusions.

- (65) a. DP+DP is transformationally derived from DP+PP.  
– based on UTAH

<sup>17</sup> This no doubt should be related to the fact that in the extant descendants of Latin, the double object construction does not appear.

- 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 Heavy NP Shift
- d. VP Shell for every VP with two things in it.
  - scope facts
  - new account of Heavy NP Shift
  - Gapping facts
  - idiomatic  $\theta$ -role to direct object

Kayne’s account of the complementation patterns in nominals speaks on behalf of Larson’s VP shell architecture for double object constructions, but is inconsistent with his proposal that this architecture be used for all two complement constructions. There is a problem, then, lurking among these conclusions. Indeed, there are several.

### 6.3 Complement structure and Object Shift

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  $\theta$ -roles these two structures license. Consider, by way of illustration, the examples below.

- (66) 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.
- (67) a. John bought a book for Bill.
- b. John bought Bill a book.
- c. John taught French to Mary.
- d. John taught Mary French.

## 6. Complement Structure

---

- 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 (67). 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 (66) lead to the same conclusion. It looks like the external  $\theta$ -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  $\theta$ -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 HEAVY NP SHIFT story. It seems unable to distinguish the following two sentences, as Jeff Runner points out.<sup>18</sup>

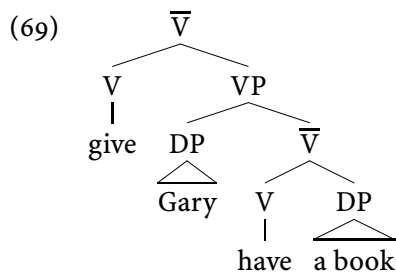
- (68)
- a. I gave a call yesterday to all my friends in Duluth.
  - b. \* I gave yesterday a book to all my friends in Duluth.
  - c. \* I gave yesterday Mary a book.

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 HEAVY 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 HEAVY

<sup>18</sup> See Runner (1995).

NP SHIFT is faulty. Until we can see a successful treatment of HEAVY 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 (69), in which the *have* verb is unspoken.



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 (69). "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 (69). 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

## 6. Complement Structure

---

the complements; consider:

- (70) a. Mary explained [<sub>DP</sub> the brats] [<sub>PP</sub> to their parents].  
b. ?? Mary explained [<sub>PP</sub> to their parents] [<sub>DP</sub> the brats].  
c. ?? Mary explained [<sub>CP</sub> that the brats needed help] [<sub>PP</sub> to their parents].  
d. Mary explained [<sub>PP</sub> to their parents] [<sub>CP</sub> that the brats needed help].

Here the  $\theta$ -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  $\theta$ -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  $\theta$ -roles. The  $\theta$ -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  $\theta$ -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 (71) illustrates.

- (71) a. Mary was shown a book.  
b. \* A book was shown Mary.

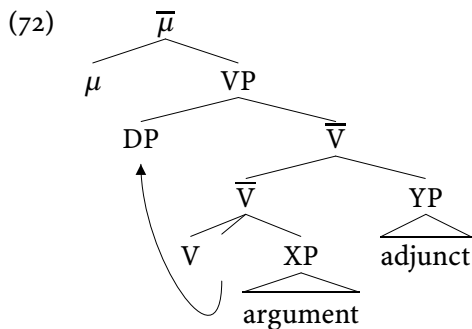
This suggests that it is only the first object which bears "structural" Case, which is the kind of Case that Passive robs. Let's define structural Case as the Case assigned by the functional head above VP.<sup>19</sup>

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  $\mu$ . In Johnson (1991) I tried to draw a connection between these phenomena in English and a process seen in the North Germanic languages in which Accusative Case-marked DPs overtly move leftwards out of the VP. We briefly saw the effects of this process in connection with our discussion of floated quantifiers in Chapter 3, where we saw that objects in Icelandic could shift leftwards out of the VP, leaving a quantifier behind. This process is known as "Object Shift" in the North Germanic literature; see Holmberg (1986).

<sup>19</sup> This is just a strengthening of Larson's definition of structural Case. He proposed that a verb assigned structural Case only when it is governed by the functional head governing VP — this was  $I^0$  for him.



If  $\mu$  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  $\mu$ . 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 (72) to cases where a verb is followed by two complements.



In (72), the complement DP has moved into a position that is governed, hence Case-marked, by  $\mu$ . 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 (72), 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.

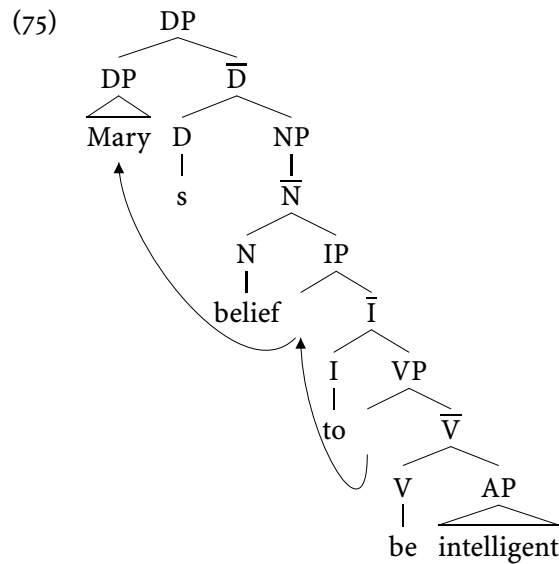
Further, because only DPs are subject to the Case Filter,<sup>20</sup> we account for the fact that it's DPs, irrespective of the  $\theta$ -roles involved, that precede other complements. Note also that once the DP has moved into Specifier of VP, there is a constituent that includes the verb and secondary 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

<sup>20</sup> Note, then, that we will have to abandon Stowell's proposal that CPs are also subject to the Case filter.

## 6. Complement Structure

---



argument suffered from not giving (an entirely satisfactory) account for the absence of cases like the following (73).

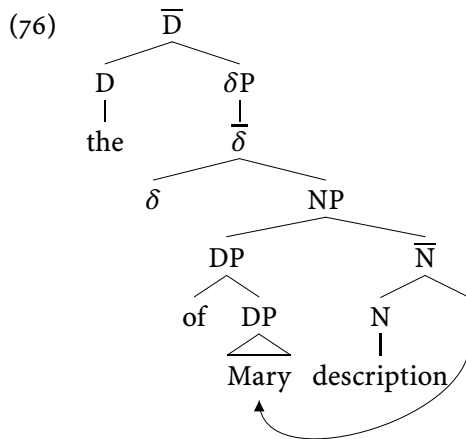
- (73)
- a. \* John's belief of Mary to have left
  - b. \* the editor's assumption of the article to contain several errors
  - c. \* John's belief (of) Mary a genius
  - d. \* their assumption of John dangerous
  - e. \* her teacher's gift of Mary of the letter
  - f. \* her rental of John of office 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 (74) 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 (75) above is blocked.

- (74) \*Mary's belief to be intelligent

The ungrammaticality of (73) might be seen as having a similar source. Suppose, for concreteness, that *of* is a Case prefix, not a preposition. 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  $\mu$ , then we will want to do the same for the *of*-marker within NPs. Let's call the head that assigns *of*,  $\delta$ . This would give a nominal like *the description of Mary* a representation something like (76) below. This isn't quite right, of course,

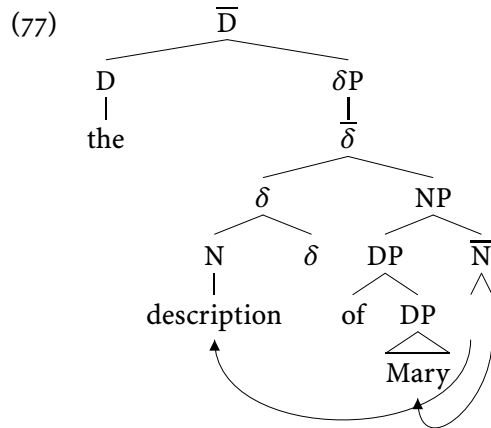


since it puts the *of* marked DP to the left of the noun that it is an argument of. In fact, that is a problem with what we have done with accusative Case marked complements too — the parse in (72) wrongly puts the object to the left of the verb. We're about to address this problem for the situation involving accusative Case, but for the situation in (76) let's adopt the solution that allows the noun to move to  $\delta$ . This, as we shall see, is one solution to the parallel problem in (72). This would transform (76) into (77) on the following page. This gets the word-order right.

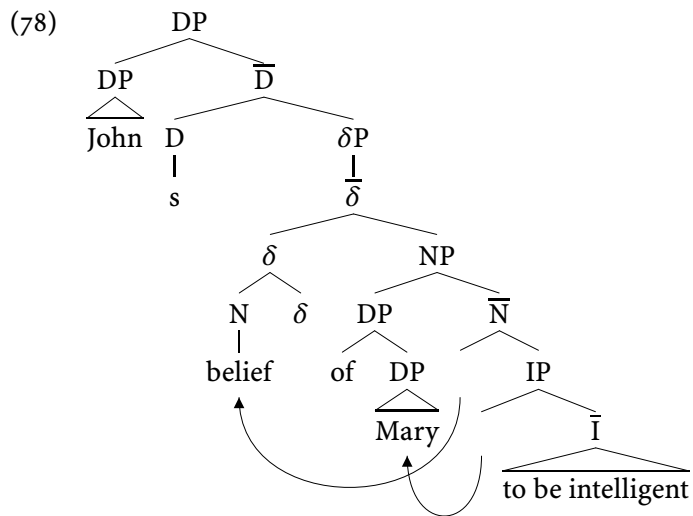
Finally we're prepared to consider what the source of the ungrammaticality of the examples in (73) might be. The relevant property of these examples is that the DP argument that is in need of the *of*-Case is embedded within a clausal complement to the noun. If the suggestions above are correct, this DP will need to move into proximity to  $\delta$  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

## 6. Complement Structure

---



same reason that (74) is. Just as in (75), the DP in (73) will have to illegally move out of the clause to get Case, as (78) on page ?? indicates. The general idea here, then,



is to preserve the parallelism between DPs and CPs/IPs that the analysis of DPs 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 (73).

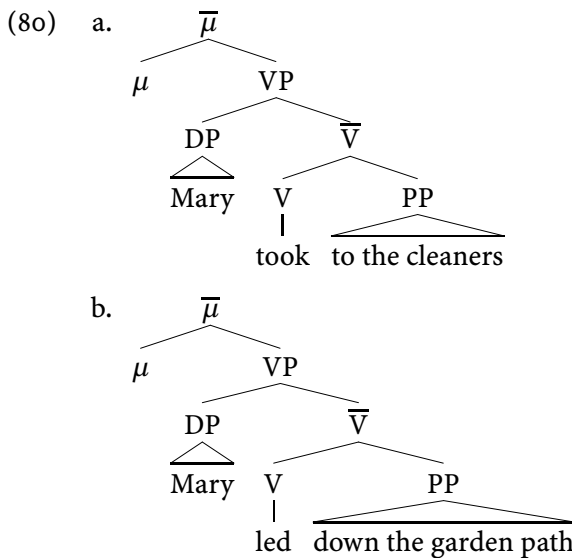
There are several reasons for adopting the thesis that accusative Case marked complements move leftwards to get Case in English, then. If this thesis is embedded in a structure like (72), we preserve Larson’s account for the scope facts, and for his Gapping data. But we are also able to preserve Stowell’s 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 (72), however. And that is his account for the idiom facts. Recall that he pointed to examples such as (79) and concluded that the first object must be able to get its interpretation – perhaps its  $\theta$ -role – by combining with a phrase made up of the verb and the second object.

- (79) a. Jerry took Mary to the cleaners.  
 b. John led Mary down the garden path.

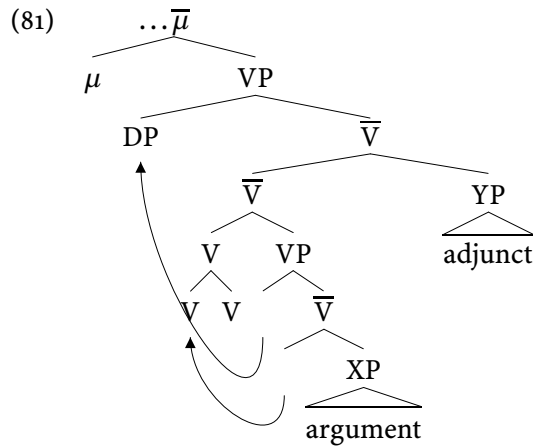
In examples such as these, the phrases *took to the cleaners* and *led down the garden path* get an idiomatic, fixed, interpretation. And, moreover, these expressions seem to take as one of their arguments *Mary*. If we hold constant our assumption that sisterhood is the structural relation which  $\theta$ -roles are assigned, then this would seem to force us to see D-structure parses such as (80) in these examples.



Because we are taking  $\theta$ -roles to be assigned at D-structure, we will want to let these examples have this sort of organization at D-structure. (72) doesn't do that. As a response to this problem, we could import some of Larson's structure into (72), maybe along the lines of (81).

## 6. Complement Structure

---



This merely embeds Larsonian VP shells inside the functional projection that assigns Accusative Case,  $\mu P$ .

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 (82) and (83).

- (82) a. Jerry explained to Sally that syntax is gripping stuff.  
 b. ?? Jerry explained that syntax is gripping stuff to Sally.
- (83) 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 (82) 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 (84).

- (84) 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.

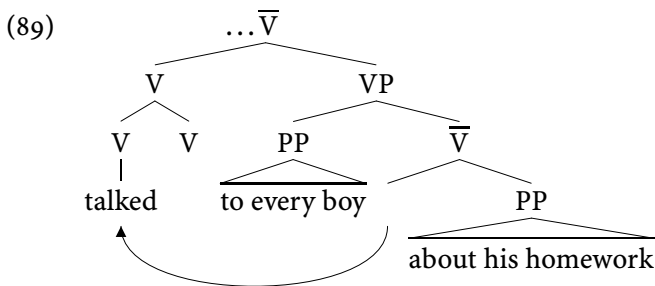
Now let us consider the scope facts which emerge with these terms. When a preposition phrase comes before a clause, it appears to have scope over that clause. The same thing appears to be true with respect to a PP that precedes another PP.

- (85) a. She explained to every boy<sub>1</sub> what he<sub>1</sub> would be fed.  
 b. She carefully explained to the kids<sub>1</sub> when each other<sub>1</sub>'s parents would pick them up.  
 c. \* She explained to the poor bastard<sub>1</sub> when Sam<sub>1</sub> would be paid.
- (86) a. She talked to every boy<sub>1</sub> about his<sub>1</sub> homework.  
 b. She talked to the boys<sub>1</sub> about each other<sub>1</sub>'s homework.  
 c. \* She talked to the poor bastard<sub>1</sub> about Sam<sub>1</sub>'s homework.

But this is not true when the *to* phrase follows the other complement, as in the following cases.

- (87) a. ? She explained what he<sub>1</sub> would be fed to every boy<sub>1</sub>.  
 b. \* She carefully explained when each other<sub>1</sub>'s parents would arrive to the kids<sub>1</sub>.  
 c. She explained when Sam<sub>1</sub> would be paid to the poor bastard<sub>1</sub>.
- (88) a. ? She talked about his<sub>1</sub> homework to every boy<sub>1</sub>.  
 b. \* She talked about each other<sub>1</sub>'s homework to the boys<sub>1</sub>.  
 c. She talked about Sam<sub>1</sub>'s homework to the poor bastard<sub>1</sub>.

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 (81), where the first phrase has the second in its scope. For example, in the PP+PP scenario, we might imagine a parse like (89).



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  $\theta$ -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 (90).

## 6. Complement Structure

---

(90) THE PROJECTION PRINCIPLE

- a. Let  $\alpha$  have c-selection requirements, and  $\alpha^{n=1,2,\dots}$  be the set of  $\bar{X}$  projections of  $\alpha$ . Then there must be an  $\alpha^x$  such that every phrase  $\beta$  within  $\alpha^x$ ,  $\beta \notin \alpha^n$ , is c-selected by  $\alpha$ .
- b. If  $\alpha$   $\theta$ -marks  $\beta$ , then  $\alpha$  and  $\beta$  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,  $\mu$ , 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 permeable 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  $\theta$ -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  $\theta$ -role is assigned to Specifier of VP (or other XP, as appropriate).

Another conflict with our earlier conclusions that crediting  $\mu$ , 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. 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.

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

Note, however, that English patterns with French when the verb's complement is not a DP, as in (92).

- (92) 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 HEAVY NP SHIFT, however this process is to be described, because it leaves the categorial contrast below puzzling.

- (93) 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 (91). 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 (94).<sup>21</sup>

- (94) Mary [<sub>V</sub> handed out] the toys.

<sup>21</sup> These data are parallel, then, to the separable prefix verbs we examined in homework.

## 6. Complement Structure

---

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.

- (95) 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

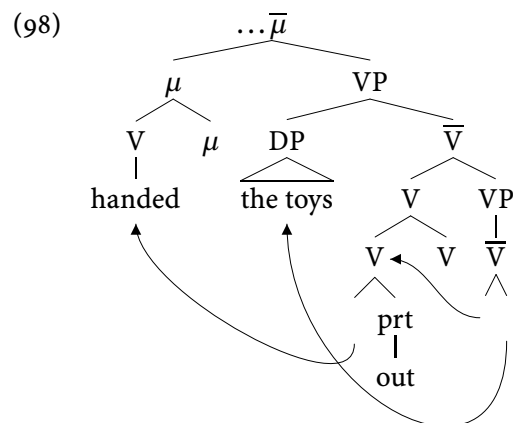
But, interestingly, it looks like there is room between the verb and particle, as shown by cases like (96).

- (96) Mary handed the toys out.

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 [<sub>v</sub> 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.

- (97) 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 (96) is as in (98).

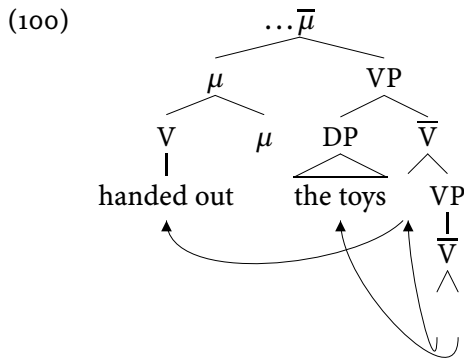


Because only DPs move to the Specifier of the higher VP, only DPs will be able to precede the particle, thereby deriving the contrasts in (97).

It's also possible for the particle to precede an object DP, however, as in (99).

(99) 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 (100).



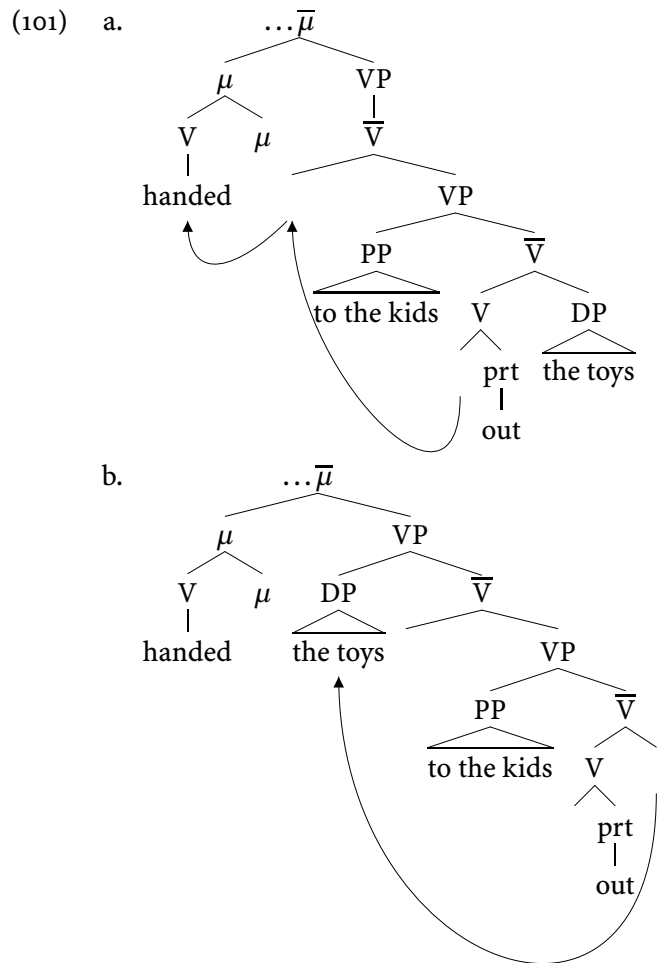
Particle constructions, then, also suggest that main verbs in English do in fact move past the position that object DPs are forced to surface in. So let us embrace this as our method of solving the problem that moving object DPs into Specifier of VP gives rise to – in particular, let us suppose that 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 (89) 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 (101) on the following page. 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.

There are a couple of directions that we might go in seeking a solution to this problem. One would be to reverse Larson's decision about which of his VP Shells holds the overt, lexical, verb. If we assign to the higher of these V<sup>0</sup>s this role, then (101) would instead have the representation in (102) on page 223. For an approach something like this, see Pesetsky (1995).<sup>22</sup> This proposal will have trouble making

<sup>22</sup> Pesetsky does not posit an empty embedded VP, but instead suggests that it is the PP we see as the second complement that makes up the embedded small clause. The first complement would then be

## 6. Complement Structure

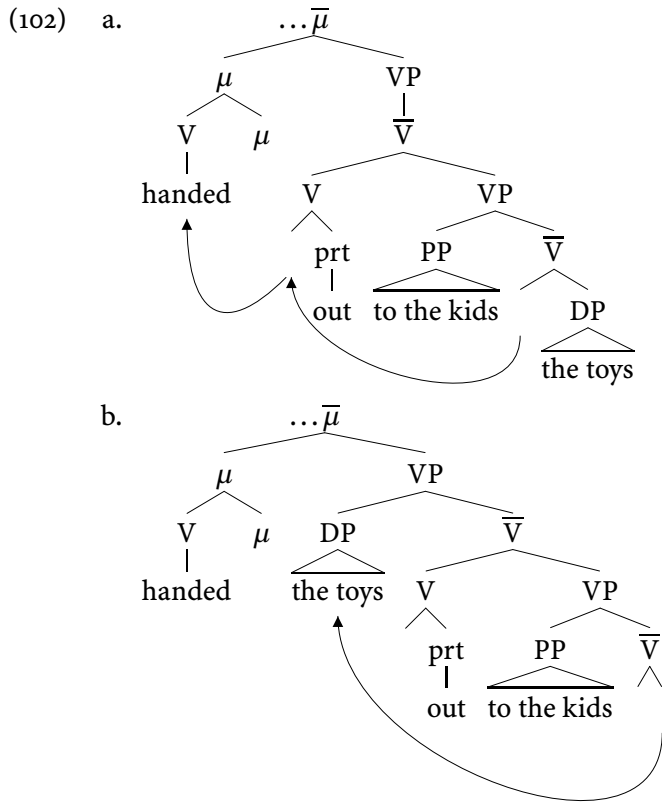


sense of Kayne's nominalization facts, however, as it claims that even verbs whose two complements are a DP and PP will have an embedded small clause in them. This would predict that both double object and these double complement constructions should be blocked in nominalization contexts. This, as we've seen, isn't the case however.<sup>23</sup>

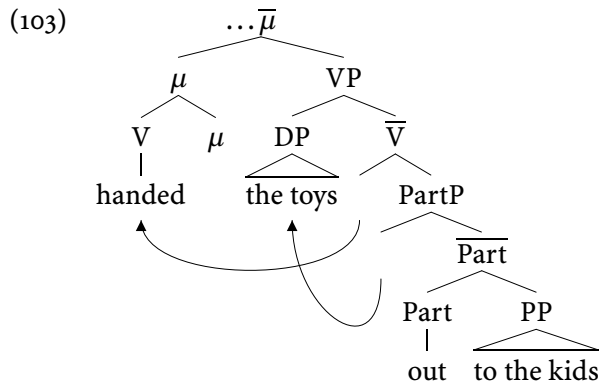
Another alternative would be to abandon our starting assumption about the particle construction: that the particle is inserted in the syntactic representation in the same position that its accompanying verb is. Indeed, one popular account of the particle construction claims that the particle heads a small clause that is embedded

within the Specifier of this PP.

<sup>23</sup> Pesetsky consequently reanalyzes Kayne's paradigm.



as a complement to the verb, as indicated in (103).<sup>24</sup>



As can be seen, it's crucial on this account that something guarantee that the “subject” of the particle phrase be the DP and not the PP in such cases. This is one of the difficulties faced by this account.

<sup>24</sup> See Kayne (1984b) and Dikken (1992).

## 6. Complement Structure

---

Rather than choose between these alternatives, or consider others, I will simply leave this problem unexamined. I suspect a resolution of the problem will require a better understanding of how the component parts of the VPs we are looking at get interpreted; more particularly, it will require a better understanding of what the meaning of the hidden small clauses we've been hypothesizing are. We'll take up this issue to a small degree in the chapter that follows.

I took the view that verbs in English move farther than just  $\mu$  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.

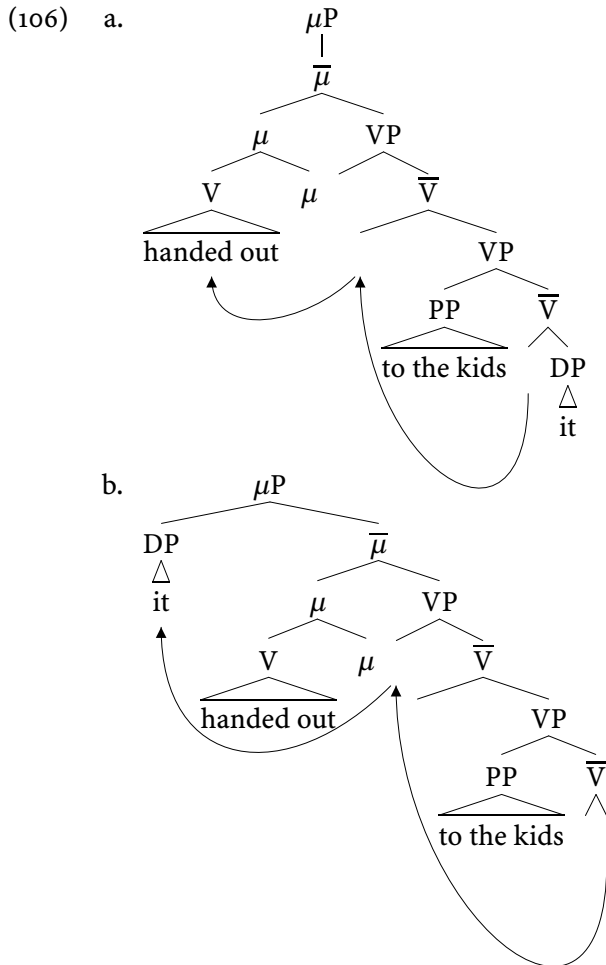
- (104) a. \* Sally handed out it.  
b. Sally handed it out.

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.<sup>25</sup>

- (105) 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 (105) 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  $\mu P$ , then we can explain why they must precede the particle. But this will require that verbs move beyond  $\mu P$ , because they always come to the left of objects, even pronominal ones. So, (104b), for instance, will receive a representation like that in (106b).

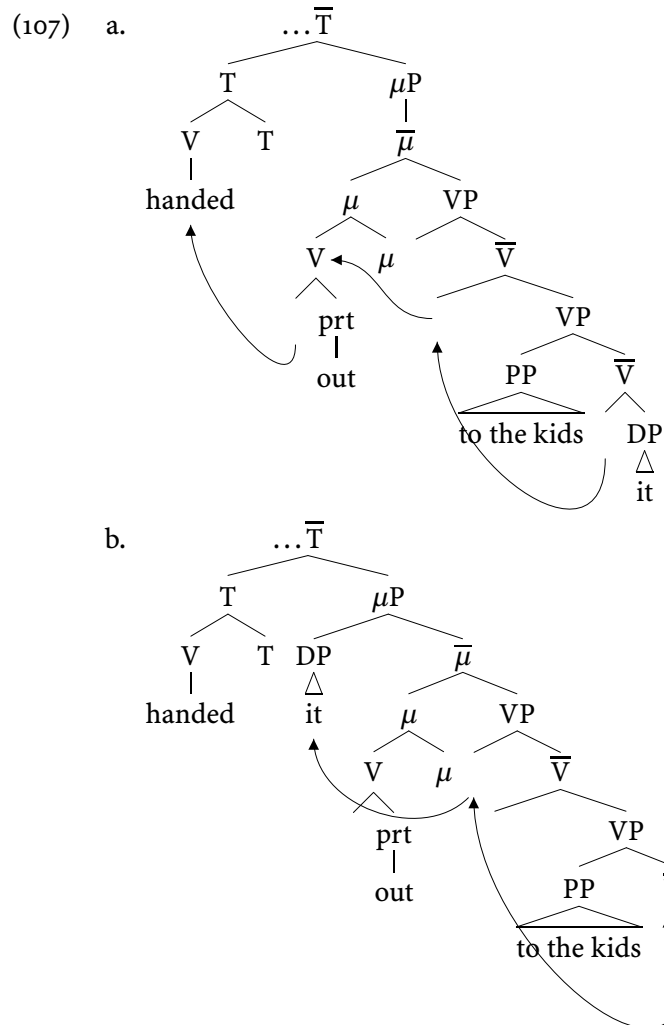
<sup>25</sup> See Diesing (1992) for a discussion of these examples.



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, (106) will produce a surface representation like that in (107) on the following page. 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

## 6. Complement Structure



reside on the interior verbal part. This is what the alternations in (108) indicate.<sup>26</sup>

- (108) 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 check-

<sup>26</sup> An exception to this pattern seems to be *able* suffixation, which prefers to attach outside the entire verbal complex: *look up-able* vs. \**lookable up*.



ing 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.<sup>27</sup>

<sup>27</sup> In 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.



# 7

## 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.

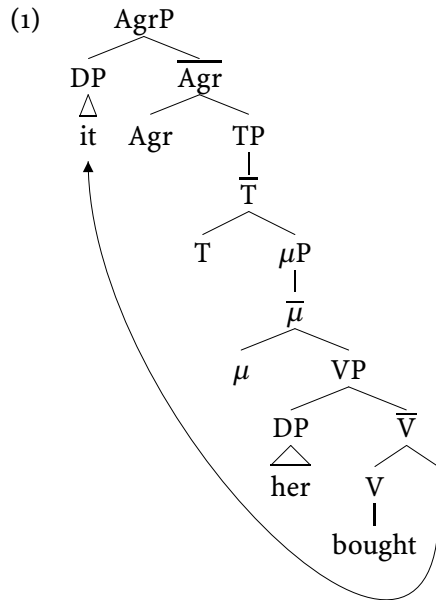
But, as touched on, this account runs into several incompatibilities with conclusions we have reached earlier. If the mechanism that is responsible for making “object” DPs move is that accusative Case is assigned by  $\mu$ , and not verbs, then verbs are not distinguishable by virtue of their ability 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 driven into Specifier of VP to be Case marked by  $\mu$ , then where are subjects placed — up to now we have taken the view that external  $\theta$ -roles are assigned to subjects in this position. Could Specifier of VP be occupied by both subject and object?

### 7.1 Getting into the right position

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

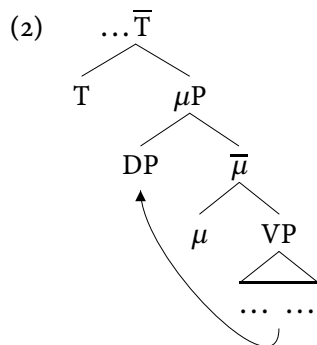
## 7. Subjects and Complex Predicates

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) below. 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  $\mu P$  always, and not into the Specifier of VP that  $\mu$  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  $\mu$  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  $\mu$  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  $\mu P$ , under (3) they will have to move to get to Specifier of  $\mu P$  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. Specifier of DP is assigned Genitive Case.
  - c.  $X^0$  assigns its Case to the position  $\alpha$  only if  $X^0$  governs  $\alpha$ .

If the structural Cases are Nominative, Genitive and Accusative, then Accusative is the only structural Case assigned under government.<sup>1</sup> 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 (6). The other instances of accusative Case assignment could be instances of inherent Case. Every other situation involving accusative Case arises when the Case assigner also assigns  $\theta$ -role to the relevant DP. For instance, prepositions assign their accusative Case to the DP that is their argument, as in (5).

- (5) Jerry fell [<sub>PP</sub> under [<sub>DP</sub> the tree]].

The complementizer *for*, however, clearly assigns a Case to a DP that is not its argument.

- (6) I prefer [<sub>CP</sub> for [<sub>AgrP</sub> Mary to be on time]].

<sup>1</sup> Recall that government is the relation defined to express when a Case assigner and Case bearer are sufficiently *klose* to each other.  $\alpha$  GOVERNS  $\beta$  iff  $\alpha$  c-commands  $\beta$  and there is no more than one phrase that contains  $\beta$  but not  $\alpha$ .

## 7. Subjects and Complex Predicates

---

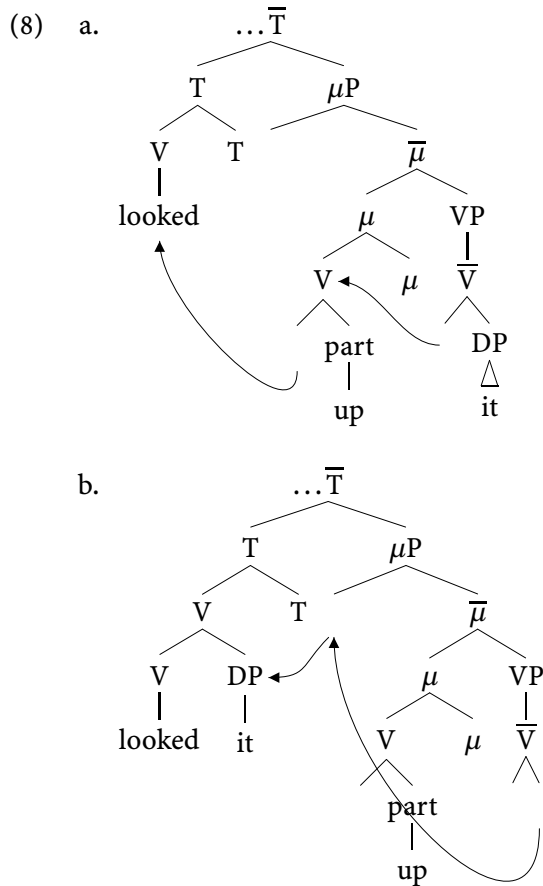
However, some have suggested that this case has a different analysis — we'll examine one of these later in this chapter — and this clears the way towards a more uniform treatment of structural Case assignment; something along the lines of (7), for instance.

- (7) Structural Case is assigned by  $\alpha$  to its Specifier position
- a. finite Agr<sup>0</sup> assigns Nominative.
  - b.  $\mu^0$  assigns Accusative.

What's lost 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  $\mu$ 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  $\mu$ P, and non-pronouns are in Specifier of  $\mu$ P, this means that pronouns must have a position that is higher than Specifier of  $\mu$ 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 them. If we let particles move with the verb out of  $\mu$ P, say as far as T, this would give us representations like (8) on the facing page. Note that in (8) the particle has been stranded in  $\mu$ . 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<sup>0</sup>.

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  $\mu$ P, rather than the other way round.

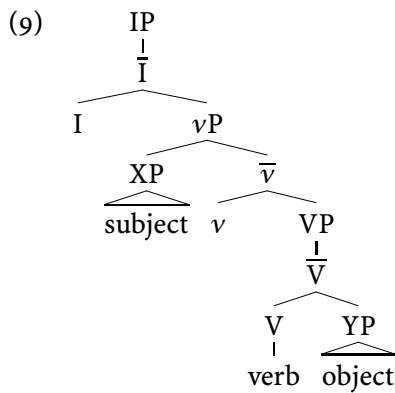
## 7.2 Subject Arguments

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

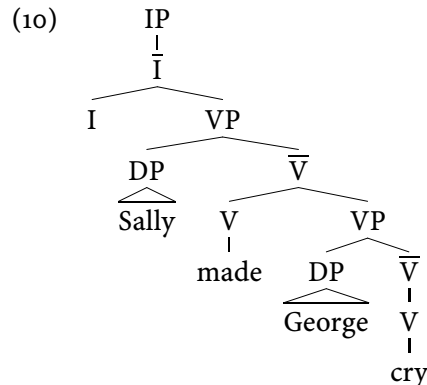
## 7. Subjects and Complex Predicates

of are connected in such a way that they come together, giving the appearance, in normal circumstances, of belonging to the same predicate. This approach would strip from the meaning of verbs that component responsible for assigning a subject  $\theta$ -role, assign this role to another hidden predicate in the sentence. This thesis is rather like the one we entertained for double object constructions, then, in which a similar deconstruction was posited.

The usual way of representing this thesis is as in (9), where  $v$  represents the hidden predicate that assigns the subject  $\theta$ -role. (I have ignored in this representation the existence of  $\mu P$ , and neglected to expand “IP” into its component TP and AgrP.)

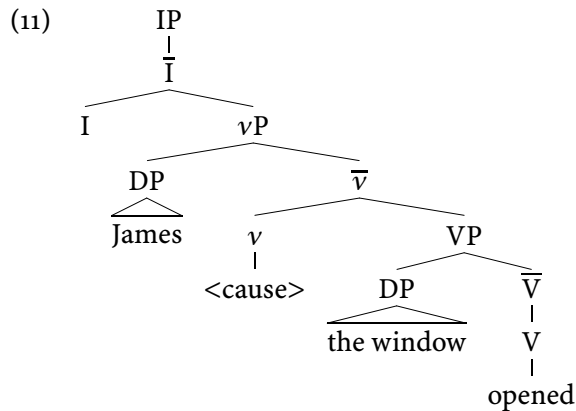


Different authors express the relationship between  $v$  and VP, and the consequent meanings of  $v$  and the verb, differently. On one view, which goes back to ideas of the Generative Semanticists, the VP is an object of  $v$ ; this would fit (9) to the requirements of the Projection Principle which requires that the sisters to heads be arguments of those heads. One model for this view are causative constructions like those in (10).



If  $v$  has a meaning along the lines that we associate with the verbs *make* or *cause*, then a sentence like *James opened the window* might have the representation in (11).





Another suggestion makes use of a semantics that credits verbs with a meaning in which they describe properties of “events.” This proposal originates with Davidson (1967), and an application of this idea to  $v$  is found in Kratzer (1996). To see how this idea works, let’s make a short digression into the semantics of argument structure.

### 7.2.1 Argument Structure

We have been using the language of  $\theta$ -roles to describe the relation between verbs and their arguments, but without any particular commitment to the content of these labels. Thus, for instance, to say that *open* assigns the THEME  $\theta$ -role to *the window* in (11) has meant nothing more than to say that *the window* is an argument of *open*. The “THEME” part isn’t meaningful, on this use of the terminology. We could just as well think in more general terms: the verb is a function that takes its object as argument and gives as a result a meaning that is associated with the VP. In the case of the VP in (11), we can surmise that its meaning should be roughly equivalent to that of the sentence in (12).

(12) The window opened.

A common way of thinking about the meaning of sentences is that they describe the conditions under which what they say would be true. If we ignore the semantic contribution that tense makes in (12), these conditions could be rendered somewhat disfluently as (25).

(13) (12) true if the object referred to by *the window* gets into the state referred to by *open*.

This, of course, is not very revelatory. It describes the meaning of *the window opened* by making cryptic reference to the meanings of *the window* and *open*. What

## 7. Subjects and Complex Predicates

---

would be helpful is to know how to determine what “the object referred to by *the window*” and “the state referred to by *open*” are. If we had that, then we could take (25) to be real step towards defining the meaning of (12). Since we will only be concerned here with the methods by which arguments and argument takers combine semantically, however, it will be sufficient to gloss over this matter. We can do with informal descriptions of the meanings of arguments and their argument takers and concentrate just on how those meanings are brought together.

If (25) is also roughly the meaning of the VP in (11), then we can deduce what the meaning of the verb is under the plausible assumption that the meaning of this VP is result of combining the meanings of its internal constituents. Because there is nothing that semantically distinguishes the verb *open* from the  $\bar{V}$  it heads, let's assume that the meaning of the  $\bar{V}$  is the same as that for *open*. Thus the meaning of *open* should be something like (14).

(14) true if  $x$  gets into the state referred to by *open*.

The  $x$  in this formula is a variable, whose value will be determined by the object's meaning. We can think of the meaning of a verb, then, as a function which applies to the meaning of its arguments and returns a meaning that corresponds to roughly that of the sentence.

If a verb combines with more than one argument this picture will have to be embellished with something that orders the application of the verb's meaning to each of the arguments. For instance, suppose, keeping with the thesis that transitive clauses have  $v$  in them, that the meaning of the VP in (15) is what (16) paraphrases.

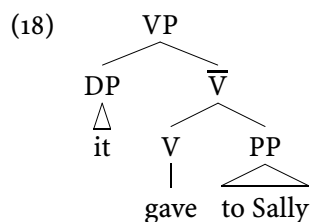
(15) She [<sub>VP</sub> gave it to Sally].

(16) meaning of VP in (15) = true if the referent of *it* is in the relation named by *give* to the individual referred to by *Sally*.

The meaning of the verb *give*, on this view, would be (17).

(17) true if  $x$  is in the relation named by *give* to  $y$ .

There are two open variables in (17), one for each object. To ensure that the variables get associated with the correct objects, we'll need a way of ordering how the function in (17) applies to the arguments. The D-structure parse for the VP in (15) is as shown in (18).



The principle that requires “ $\theta$ -roles” to be assigned under sisterhood can be interpreted on this way of talking as requiring that the application of the verb’s meaning be ordered by the syntactic sisterhood relation. That is, we can translate the sisterhood restriction on  $\theta$ -role assignment to something like (19).

- (19) The meaning of  $\alpha$ , a phrase, is the result of semantically combining  $\alpha$ ’s daughters.

This is more general than the requirement that  $\theta$ -roles be assigned under sisterhood; it spreads this sisterhood requirement onto all semantic relations. Let’s gamble that this won’t introduce trouble down the road. In the case of (18), it will require that the meaning of *gave* apply first to *to Sally* to form a meaning that will then apply to *it*. Our D-structure representations in concert with (19) will order, then, how the meaning of verb is combined with the meanings of its arguments.

What’s needed now is a way of translating that ordering into procedure that puts the arguments’ meanings into the right variable positions. For (18), for instance, we need a way of ensuring that the fact that *gave* combines semantically first with *to Sally* means that the meaning of *to Sally* fills the  $y$  position in the formula that is *gave*’s meaning. A common way of doing this is to use Alonzo Church’s  $\lambda$ -operator. The  $\lambda$ -operator provides a very general method of defining functions of the sort that we are associating with the meanings of a verb. The  $\lambda$ -operator can be defined as in (20).

- (20) Let  $P(a)$  represent an expression  $P$  that contains a term  $a$ .  $\lambda x.P(x)$  is a function that when applied to  $a$  produces  $P(a)$ .

The meaning of *gave*, then, could be expressed using  $\lambda$ -operators as (21).

- (21) the meaning of *gave* =  $\lambda y.\lambda x.$ true if  $x$  is in the relation named by *give* to  $y$ .

The order of  $\lambda$ -operators will steer how the variables within the formula are replaced.

Let’s consider how this will work in the case of (18). Suppose that the meaning of *to Sally* is something like “the individual referred to by *Sally*.” Combining the meaning of *gave* in (21) with this will yield (22).

- (22) the meaning of  $\bar{V}$  in (18) =  $\lambda x.$ true if  $x$  is in the relation named by *give* to the individual referred to by *Sally*.

The meaning of the VP in (18) will then be derived by combining this meaning with the meaning of *it*. If the meaning of *it* is something like “the individual referred to by *it*,” then the meaning of (18) will be (23).

- (23) the meaning of the VP in (18) = true if the individual referred to by *it* is in the relation named by *give* to the individual referred to by *Sally*.

## 7. Subjects and Complex Predicates

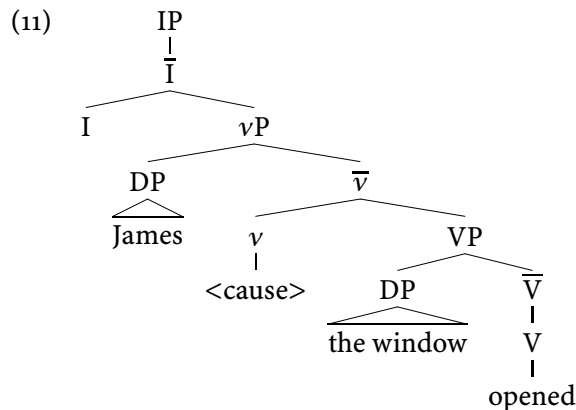
---

This, then, might be a way of filling in what it means to “assign a  $\theta$ -role”; it amounts to having an argument fill an open slot in the meaning of a verb. As can be seen, there is no content to the talk about  $\theta$ -roles on this view. To say that a verb “assigns a Goal  $\theta$ -role to  $\alpha$ ” amounts to saying that  $\alpha$  fits into a certain position in the verb’s meaning. On such a view, principles or generalizations that make reference to  $\theta$ -roles, like UTAH or Larson’s use of a  $\theta$  hierarchy to determine how arguments are placed in D-structures, aren’t sensible. Their effects would have to be expressed in some other way. This view of the predicate-argument relation is what Dowty (1989) calls the “ordered argument” view.

On the ordered argument view of argument structure,  $v$  would be merely another two-place function, relating its subject argument to the meaning of the VP that follows. For example, we might imagine that the  $v$  in *James opened the window* has the meaning in (24).

(24) meaning of  $v = \lambda x.\lambda y.$ true if  $y$  did something that causes  $x$ .

Recall that the D-structure representation for this sentence is (11), and that the meaning of the VP headed by *opened* is (25).



(25) true if the object referred to by *the window* gets into the state referred to by *open*.

The first argument that  $v$ 's meaning will apply to is the VP, yielding (??).

(26) the meaning of  $\bar{v}$  in (11) =  $\lambda y.$ true if  $y$  did something that causes true if the object referred to by *the window* gets into the state referred to by *open*.

When this combines with *James*, the result is (27).

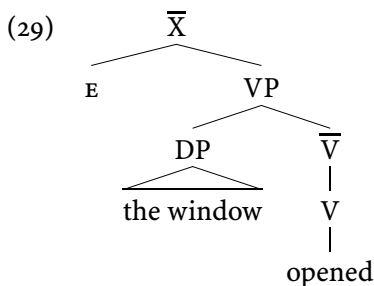
(27) true if the object referred to by *James* did something that causes true if the object referred to by *the window* gets into the state referred to by *open*.

Perhaps this is a meaning that corresponds roughly to our intuition of what (11) conveys. We'd have to have a better grip on what the expression "did something that causes true" amounts to, among other things, to be certain. But at this level of informality, it's clear that we're in the ballpark.

This, then, is one way of working out what the relationship between verb meanings and their arguments is. It expresses the thesis that verb meanings are descriptions of relations or states that hold of things. The verb *open*, for instance, describes a state its argument is in, and *give* names a relation between its two objects. There is an alternative view of verb meanings, however, in which they describe more than just properties of things. On this view, they describe relations among objects and "events," where events can be thought of as a certain arrangement of objects at some well-defined location and time. On this view, the meanings of *open* and *give* might be as in (28).

- (28) a. the meaning of *open* =  $\lambda x.\lambda z.$ true if  $x$  is in  $z$  and  $z$  is the state named by *open*.  
 b. the meaning of *give* =  $\lambda x.\lambda y.\lambda z.$ true if  $y$  is given to  $x$  in the event  $z$

This requires that sentences contain something that refers to an event, and that this term is one of the arguments of a verb. It's not clear where to place this, apparently hidden, argument, but for concreteness sake imagine that it is positioned just outside the highest VP. If we use "E" to represent the argument that refers to an event, then we might give the sentence *the window opened* a D-structure representation like that in (29).



Combining (29) with *the window* now gives to VP the meaning in (30).

- (30)  $\lambda z.$ true if the object referred to by *the window* is in  $z$  and  $z$  is the state named by *open*.

When this meaning is applied to E, the resulting meaning will be as (31) paraphrases.

- (31) true if the object referred to by *the window* is in E and E is the state named by *open*.

## 7. Subjects and Complex Predicates

---

Davidson (1967) champions the view that sentences have an event component to their meaning, and this is one way of expressing the idea.

There are some strong reasons for adopting a semantics of this sort. One, that Davidson discusses, is that it allows for a simple way of dealing with the semantics of (some) verbal modifiers. There are certain modifiers whose meanings are combined with the meaning of the phrase they modify in what appears to be a simple intersective way. For instance, the DP in (32) has a meaning that combines the meaning of the adjective and noun intersectively.

(32) the blue square

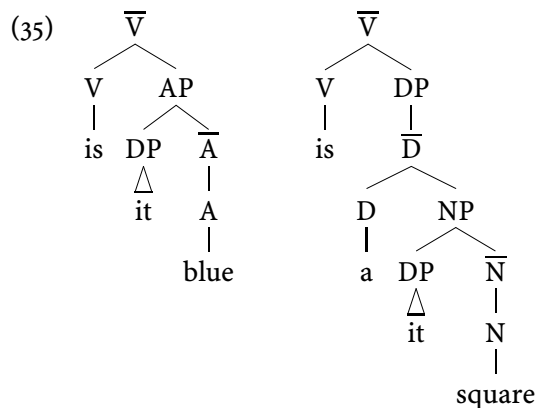
Imagine, for instance, that the meaning of *blue* is (33a) and the meaning of *square* is (33b).

- (33) a. meaning of *blue* =  $\lambda x$ .true if  $x$  is blue  
 b. meaning of *square* =  $\lambda x$ .true if  $x$  is square

That is, imagine that adjectives like *blue* and nouns like *square* are like verbs in describing the conditions under which some object has the property they name. This is suggested by the fact that nouns and adjectives can function the way verbs do in building up the meanings of a sentence. The sentences in (34), for example, have meanings along the lines paraphrased.

- (34) a. It is blue.  $\approx$  true if the object referred to by *it* is blue.  
 b. It is a square.  $\approx$  true if the object referred to by *it* is square.

These meanings arise out of the procedures we've sketched from the D-structure representations in (35) if *blue* and *square* have the meanings in (33) (and there is no meaning associated with *is* or *a*).

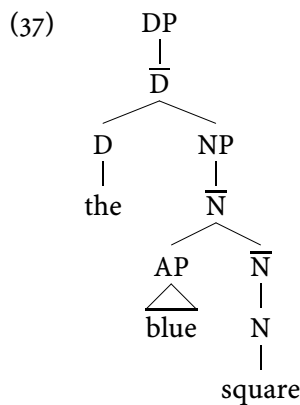


The meaning of (32) puts together the meanings of *blue* and *square* in such a way that it refers to an object that have both those properties. This requires two things.

On the one hand the DP must refer to an object, rather than describe truth conditions. And on the other it must refer to that individual by way of composing the meanings of *blue* and *square*. It seems likely that the determiner *the* is responsible for making the DP refer to an object. This can be done by letting *the* have a meaning like that in (36), where P will be the meaning that the NP it combines with has.

(36) meaning of *the* = refers to the *a* that makes P true.

So, for instance, (32), whose D-structure parse is (roughly) (37), will have the meaning in (38).



(38) refers to the *a* that makes the meaning of NP true.

To produce the right meaning for NP in (37) from the meanings of *blue* and *square* in (33) requires a rule like (39).<sup>2</sup>

(39) PREDICATE CONJUNCTION  
 If  $\lambda x.P(x)$  and  $\lambda y.Q(y)$  are sisters, then they can be combined to form  $\lambda x.P(x) \wedge Q(x)$ .

This will give the NP in (37) the meaning in (40), from which the meaning for the entire DP in (41) is achieved.

(40)  $\lambda x.\text{true if } x \text{ is blue} \wedge \text{true if } x \text{ is square}$

(41) refers to the *a* that makes [*x* is blue  $\wedge$  *x* is square] true.

What we see from this example, then, is that certain modifiers make use of PREDICATE CONJUNCTION. The point is that this same rule can be used to express the parallel instances of sentential modification, if sentences describe properties of events. For instance, imagine that *in the room* has the meaning indicated by (42).

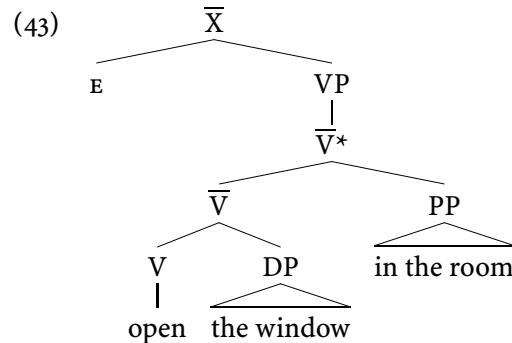
<sup>2</sup> The meaning of  $\wedge$  is roughly that of the English connective *and*.  $P \wedge Q$  is true just in case *P* is true and *Q* is true.

## 7. Subjects and Complex Predicates

---

(42) meaning of *in the room* =  $\lambda e.$ true if  $e$  is in the room

It is now a simple matter of using PREDICATE CONJUNCTION to derive the meaning of the VP in (43), under the view that verbs describe properties of events.



If (28) is the meaning for *open*, then applying that meaning to *the window* will give the resulting  $\bar{V}$  the meaning in (44).

(44) meaning of  $\bar{V}$  in (43) =  $\lambda e.$ true if the object referred to by *the window* is in  $e$  and  $e$  is the state named by *open*.

PREDICATE CONJUNCTION will now compose this meaning with (42) to give the resulting  $\bar{V}^*$  the meaning in (45).

(45)  $\lambda e.$ true if the object referred to by *the window* is in  $e$  and  $e$  is the state named by *open*  $\wedge$   $e$  is in the room.

And when this applies to  $\mathbb{E}$ , the result will be (46), which is close to what our intuitions about the meaning of this VP is.

(46) true if the object referred to by *the window* is in  $\mathbb{E}$  and  $\mathbb{E}$  is the state named by *open*  $\wedge$   $\mathbb{E}$  is in the room

Another reason for believing that verbs and the VPs they project describe properties of events is because there are some cases in which VPs are arguments which seem to refer to events. In (47), for instance, the *vP* *James open the window* transparently refers to something that Jill saw.

(47) Jill saw [<sub>vP</sub> James open the window].

The object of *saw* in this example is not a description of conditions under which James opens the window. Instead it refers to a particular situation involving James and the window, and in which they are in the “open” relation. The object of *saw* in this example is what I’ve been describing as an event.



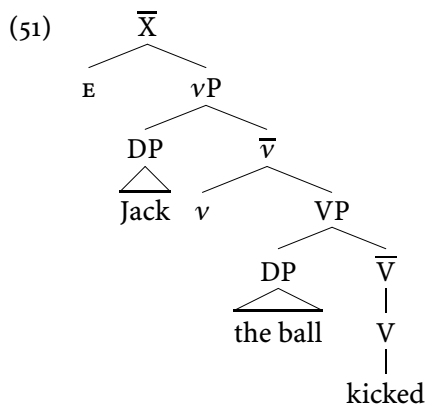
We should conclude from (47), then, that  $v$  names a property of events just as other verbs do. It's this idea that Kratzer (1996) exploits in cashing out the meaning of  $v$  and its connection to the VP that follows.<sup>3</sup> Her proposal is that  $v$  has a meaning that can be thought of as expressing something like what  $\theta$ -roles were designed for. An example involving a garden variety transitive sentence, such as (48), might have a  $v$  in it whose meaning is (49).

(48) Jack kicked the ball.

(49) meaning of  $v = \lambda x.\lambda y.\text{true if } x \text{ is the agent of the event } y.$

Consider now how this meaning will combine with the meaning of the verb phrase that follows it. This VP will have a meaning like that in (50) and, if our parses to this point have been correct, it will be fit as a complement to  $v$  as in (51).

(50)  $\lambda z.\text{true if the object referred to by } \textit{the ball} \text{ is kicked in the event } z.$



Neither VP nor  $v$  can serve sensibly as the others argument, nor can PREDICATE CONJUNCTION semantically combine them. Kratzer proposes another rule of semantic combination for this scenario that, like PREDICATE CONJUNCTION identifies the a variable in each of the formulas being brought together, and otherwise conjoins them. This rule is EVENT IDENTIFICATION.

(52) EVENT IDENTIFICATION

If  $\lambda x.\lambda z.P(x)(z)$  and  $\lambda y.P(y)$  are sisters, and  $z$  and  $y$  range over events, then they can be combined to form  $\lambda x.\lambda z.P(x)(z) \wedge P(z)$ .

EVENT IDENTIFICATION will given  $\bar{v}$  the meaning in (53), which when applied to *Jack* will produce the meaning in (54).

<sup>3</sup> In fact, she proposes that  $v$  is the Accusative case assigner,  $\mu$ . We will adopt this proposal in the next chapter. But for now it will be convenient to keep them separate.

## 7. Subjects and Complex Predicates

---

- (53)  $\lambda x.\lambda y.$ true if  $x$  is the Agent of the event  $y \wedge$  true if the object referred to by *the ball* is kicked in the event  $y$ .
- (54)  $\lambda y.$ true if the object referred to by *Jack* is the Agent of the event  $y \wedge$  true if the object referred to by *the ball* is kicked in the event  $y$ .

This meaning will apply to  $e$ , the term that refers to an event, to produce (55).

- (55) true if the object referred to by *Jack* is the Agent of the event  $E \wedge$  true if the object referred to by *the ball* is kicked in the event  $E$ .

This accords roughly to what this sentence is understood to mean.

On either of these views, solutions will have to be sought for a number of problems that arise when the subject argument is separated from the meaning of the main verb. It'll be necessary to find an account for why some verbs are not found without an accompanying  $v$  — such as *kick* — while other verbs occur freely with our without  $v$  — such as *open*. Moreover, it will also be necessary to ensure that the right  $v$  is matched with the verb. The  $v$  found in (56), for instance, should not be wrongly found in (57).

- (56) This problem frustrates me.
- (57) \* This problem kicks me.

To use our  $\theta$ -role talk, the  $v$  in (56) assigns an Experience  $\theta$ -role, and the  $v$  in (57) should assign an Agent  $\theta$ -role instead.

For this second problem, Kratzer suggests that a typology of events will do part of this work. Events might come in different kinds: states, for instance, or actions. And it might be possible to think of the different “ $\theta$ -roles” assigned to subject arguments as being a function of the sorts of events that they are subjects of. One might imagine, say, that Agents are subjects just of action events whereas Causers are subjects of state events. If EVENT IDENTIFICATION is restricted so that it only allows events of the same kind to be identified, then this will allow the  $v$  that describes a state event to combine only with verbs that similarly describe a state event. We might even hope to find an account in this for the dramatically more restricted range of subject relations when compared to object relations.

If this is the right direction, it will require a more complex method of combining  $v$  and VP in cases like *James opened the window*. In this example we want the event that *James* is the Agent of not to be identified with the (stative) event described by *opened the window*. Instead, we want the event that *James* is an Agent of to be the case of, or result in, the stative event: *opened the window*. There are many interesting problems of this type to be sorted out; this is a growth industry in the syntax/semantics arena.

7.2.2 The syntactic benefits of  $v$ 

There is another potential advantage to charging  $v$  with the assignment of subject  $\theta$ -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  $\theta$ -roles and (structural) Accusative Case assignment, repeated in (58).

## (58) BURZIO'S GENERALIZATION

If a verb assigns (structural) Accusative Case, then it assigns an external  $\theta$ -role.

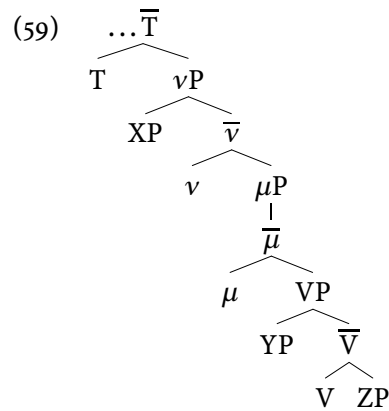
Recall that what this statement records is an unexpected gap in the logically possible outcomes, on the assumption that Accusative Case marking and the subject  $\theta$ -marking properties of a clause are independent, as the theory up to now has held. What (58) suggests, by contrast, is that Accusative Case assignment and external  $\theta$ -role assignment are not independent. If we move to the view that external  $\theta$ -roles are not assigned by verbs, but are instead assigned by  $v$ , then we can let the contingency expressed in (58) come about by way of selection. Let one of  $\mu$  or  $v$  c-select the other, and (58) will emerge. This is a proposal in Chomsky (1995b).<sup>4</sup>

The reason this approach to external  $\theta$ -roles might be helpful for our problem of delivering objects and subjects to the correct Case marked positions is that it allows the external  $\theta$ -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. This thesis would arrange  $vP$  and  $\mu P$  as indicated in (59) on the next page. On this syntax, then, Burzio's Generalization is expressed as a consequence of  $v$  selecting  $\mu P$ . Everything we have seen about the derived subjects hypothesis is consistent with (59). In particular, it correctly allows quantifiers to be floated beneath auxiliary verbs, but not below main verbs or their objects. Placing  $\mu P$  between  $vP$  and VP in this way is what Masatoshi Koizumi calls the "Split VP Hypothesis," and his 1995 MIT dissertation is the *locus classicus* for this position.

<sup>4</sup> Chomsky'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  $vP$ . Bowers' work, in turn, builds on Williams (1980).

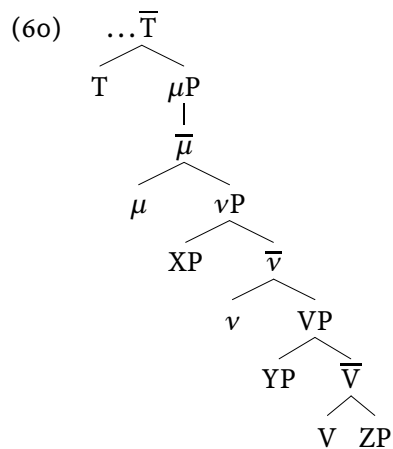
## 7. Subjects and Complex Predicates

---



### 7.3 The relative positions of $\mu$ P and $v$ P: Evidence from 'again'

Interestingly, there is some evidence that this straightforward solution isn't correct. This evidence suggests that  $v$ P is instead embedded within  $\mu$ P, as in (60) below. The



evidence comes from German data, and is based on an interesting property of some adverbs, of which *wieder* ('again') is an example.

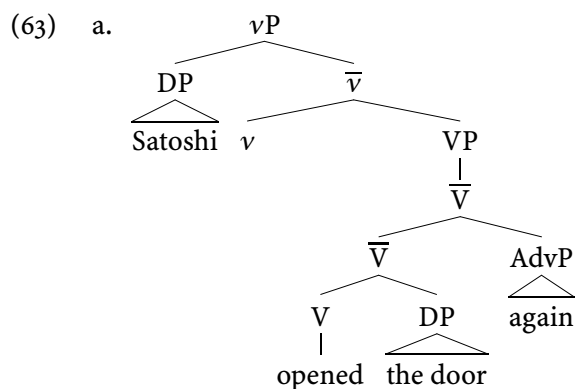
Sentences in both German and English which have this adverb in them are ambiguous in a way that suggests that *wieder/again* is capable of modifying two different things. Consider, by way of illustration, (61).

(61) 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 (61) 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 (61) has the two readings in (62).

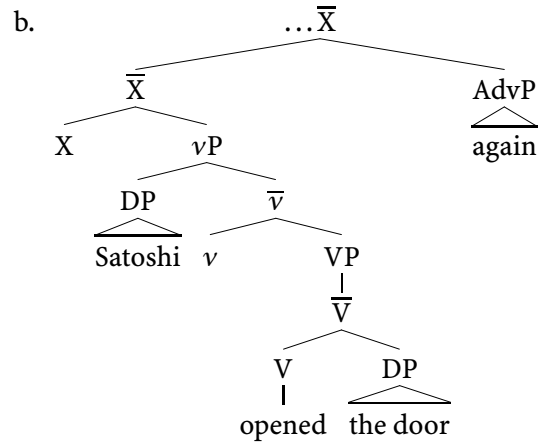
- (62) a. Satoshi opened the door, and he had done that before.  
(repetitive)
- b. Satoshi opened the door, and the door had been opened before.  
(restitutive)

von Stechow (1996) argues that this ambiguity reflects a structural ambiguity. In the repetitive reading, *again* modifies the constituent we have identified with  $\nu P$ , as in (63a). In the restitutive reading, by contrast, *again* modifies the lower predicate, as in (63b). (These parses, both below, don't reflect the movements necessary to achieve the surface representations.)



## 7. Subjects and Complex Predicates

---



If  $vP$  and  $VP$  in (63) have meanings like those described above, then (63a) has a meaning paraphrased by “Satoshi again made the door be in an open state,” and (63b) has a meaning paraphrased by “Satoshi made the door again be in an open state.”

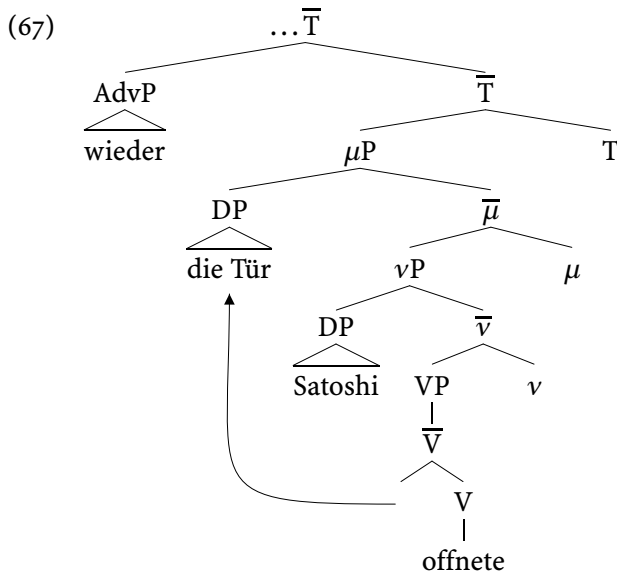
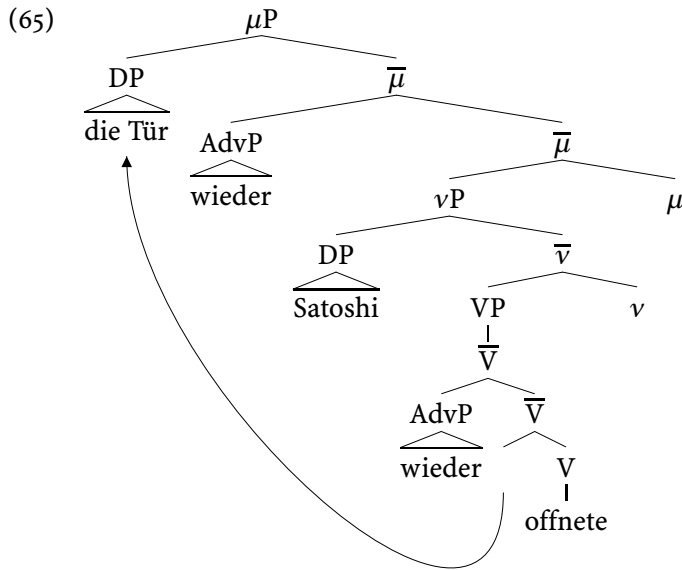
Stechow’s argument on behalf of this structural account is based on the fact that word order plays a role in disambiguating *again* in German. The German version of (61) is (64).

- (64) ...weil Satoshi die Tür wieder öffnete.  
 ...since Satoshi the door again opened.  
 ‘...since Satoshi opened the door again’

Like (61), (64) has both the repetitive and restitutive readings. On the structural account of this ambiguity that (63) sketches, this means that both  $vP$  and  $VP$  must be embedded beneath the position that the Accusative Case marked object, *die Tür*, surfaces in. If we let  $XP$  in (63b) be identified with  $\mu P$ , thereby explaining why this is the surface position of the Accusative case marked object, we could represent this ambiguity by letting *wieder* have either of the two positions indicated in (65) on the facing page. Interestingly, if *wieder* shows up to the left of the object, as in (66), only the repetitive reading is found.

- (66) ...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 (67).



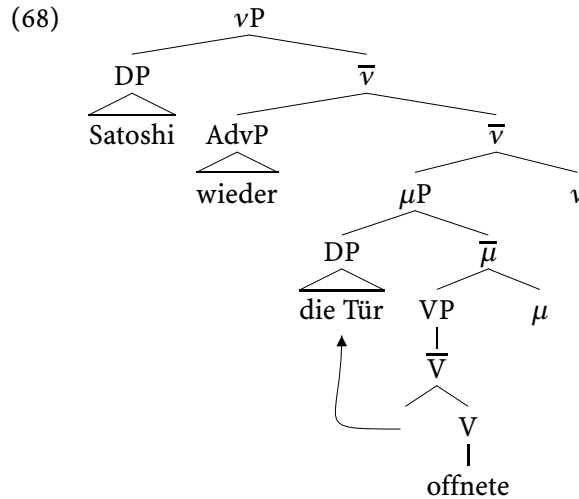
As a consequence, *wieder* will have to modify a constituent containing the “Satoshi makes” part of the meaning, and this is just the repetitive reading.

Putting  $\mu$ P above  $\nu$ P, and forcing Accusative Case-marked objects to be pronounced in Specifier of  $\mu$ P in German, gives an account, then, of the disambiguating effect that placing *wieder* before or after the object has. If  $\nu$ P could be higher than  $\mu$ P, then we should be able to give (66) the representation in (68), and this fails

## 7. Subjects and Complex Predicates

---

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 (69), then only the repetitive reading is possible.

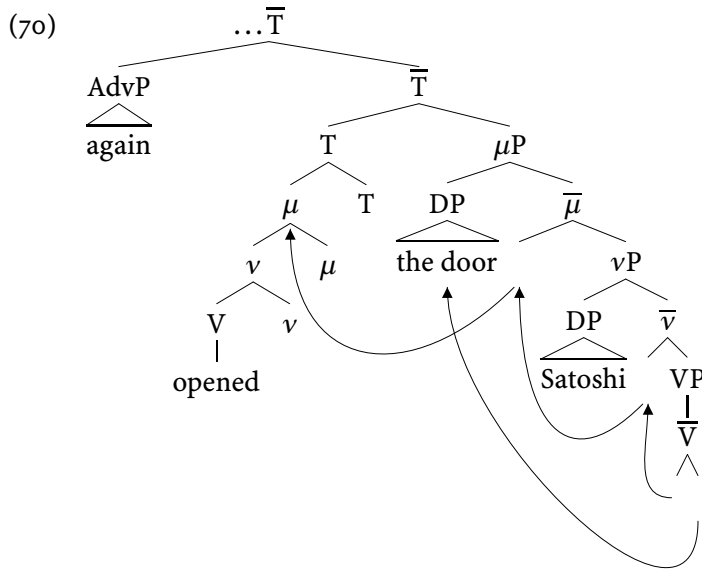
- (69) Satoshi again opened the door.  
 = Satoshi again made the door be in an opened state.  
 ≠ Satoshi made the door again be in an opened state.

Just as in the German (66), 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 (70) on the next page. As a consequence, *again* will be forced to modify a constituent that contains the “make” part of the meaning, i.e.  $vP$ , but only if  $vP$  is necessarily embedded within  $\mu P$ .

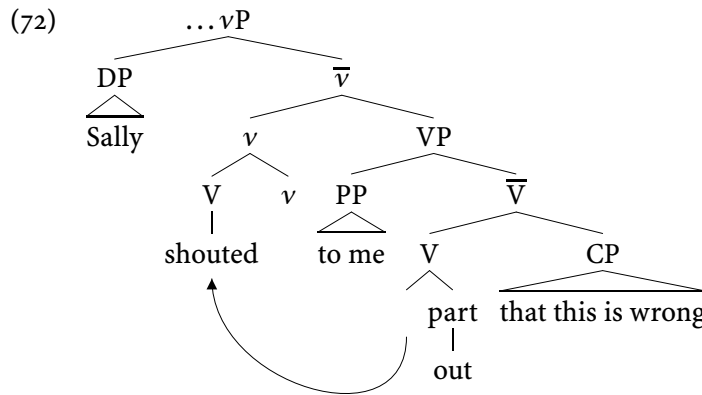
For this reason, then, let’s abandon the solution to the problem of delivering the subject and object arguments into the proper Case marked positions that is based on putting  $vP$  higher than  $\mu P$ . Instead, we should place  $vP$  within  $\mu P$ , and find another solution to the Case problem. Indeed, we could equate the highest VP of Larson’s shells with  $vP$ .

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 (71) with the representation in (72).





(71) \*Sally shouted to me out that this is wrong.  
 (compare: "Sally shouted out to me that this is wrong.")



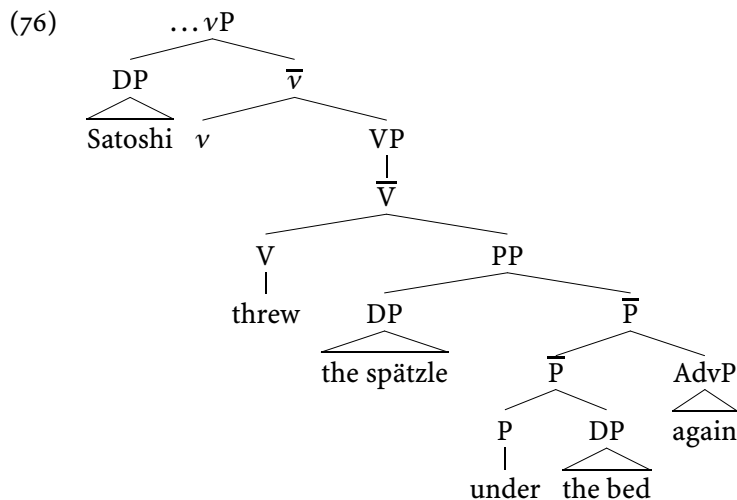
There is something, I suggested, that must be responsible for preventing particles from being stranded in their underlying position. But what could be responsible for this?

Consider now what we can learn from *again* about the structure of VPs with PPs in them. Focus on the restitutive readings in (73) through (75) (these are paraphrased for each case in the b-examples). The state that *again* says has been restored would appear to be fully described by way of the meaning of the prepositional phrase. This suggests a structure like (76), in which the PP, "the spätzle under the bed" denotes something like "the spätzle is in the state of being under the bed."

## 7. Subjects and Complex Predicates

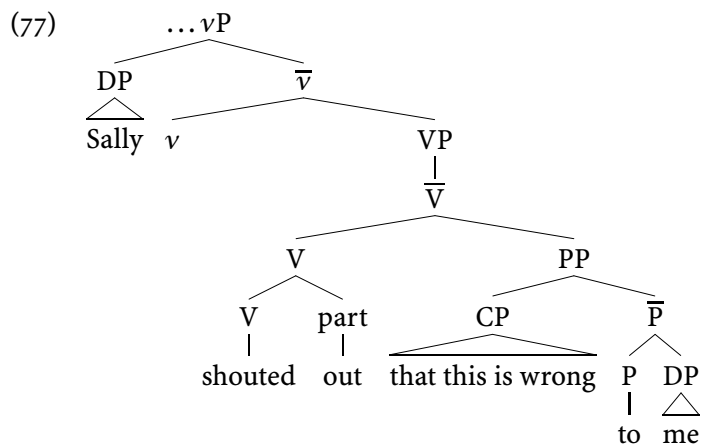
---

- (73) 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.
- (74) 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.
- (75) 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.



If this is the correct structure for Prepositional Phrases in all cases, then the problem posed by word order in (71) is solved. Indeed, Pesetsky (1995) has argued extensively, and on independent grounds, that (76) 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 (71) will be (77) on the facing page. Clearly, even if the particle is stranded in its underlying position in (77) it will remain to the left of the PP. Of course, (77) 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

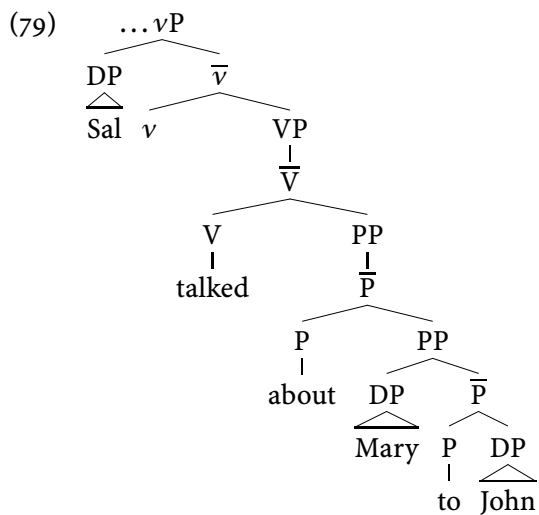


of PP is linearized so that it follows its  $\bar{P}$  when its contents are a CP.<sup>5</sup>

Perhaps the most difficult case to treat in this way are those verbs that take two PPs as complements, such as (78).

(78) Sal talked about Mary to John.

Pesetsky argues for the representation in (79) for such cases.



<sup>5</sup> What 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 (79).

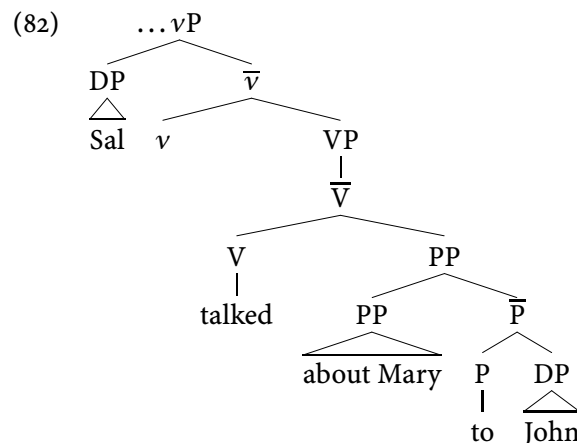
## 7. Subjects and Complex Predicates

---

One of the difficulties 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 (80) and indicate, as do the word orders in (81).

- (80) a. About John, Sal talked to Mary.  
 b. To Mary, Sal talked about John.
- (81) a. Sal talked to Mary yesterday about John.  
 b. Sal talked about John yesterday to Mary.

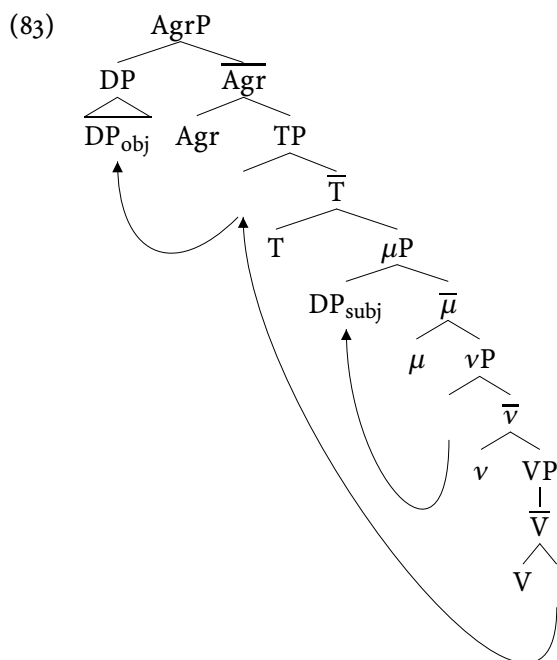
Pesetsky has a sophisticated proposal for these problems, one that would take us too far afield to explore. We might instead entertain the possibility that these cases get a representation like that in (82).



### 7.4 The Minimal Link Condition and Romance causatives

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 (83) on the next page. 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). (See also Bobaljik and Jonas (1996).)

This solution is built on the proposition that Argument movement is subject to a constraint we have not yet encountered. This constraint prevents some term from Argument 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 (84).



(84) SPECIFIED SUBJECT CONSTRAINT

If  $\alpha$  Argument moves to  $\beta$ , then there must be no  $\gamma$  in a Specifier position that c-commands  $\alpha$  but not  $\beta$ .

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 Argument 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 (85).

(85)  $[_{XP} \beta X^0 [_{YP} \gamma Y^0 \alpha ]]$

If  $\gamma$  is a subject, then it must be Case marked. This means that either YP is a CP, or  $\gamma$  is getting Case from outside YP. If YP is a CP, then movement from  $\alpha$  to  $\beta$  will violate the constraint against moving out of CPs. If  $\gamma$  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  $\alpha$  for the Nominative Case marked position. But if  $\gamma$  itself moves to the Nominative Case marked position, then it will also compete with  $\alpha$ . In brief, the

## 7. Subjects and Complex Predicates

---

Case conditions on Argument movement will derive the consequences of (84), at least over a broad range of situations.

But there are versions of Argument movement, found outside English, which are not subject to the Case conditions on Argument 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 (86), for example.

- (86) 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.'

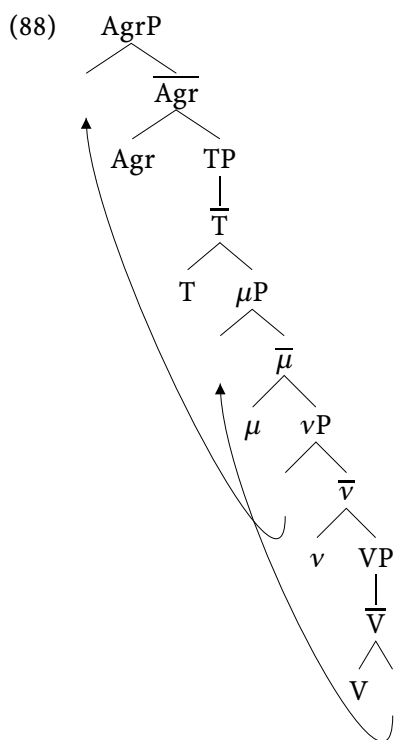
Cliticization has many of the same attributes that we have seen for Argument Movement. It obeys the Upwards Constraint and it is subject to something very like the same bounding constraints we've seen on Argument movement. In fact, from a descriptive point of view, clitics cannot move out of CPs, just 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 Argument 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.

- (87) 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 à Jean depuis longtemps  
Mary believed the gift sent to John a long time ago

In (87a), 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 (87b), 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 (83). In particular, the movement of the object into Specifier of AgrP will pass the c-commanding subject in (83), in violation of the Specified Subject Constraint. Unfortunately, this constraint also similarly blocks the grammatical outcome, shown in (88) below. Movement of the object into Specifier of  $\mu$ 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  $\mu$ P.

Chomsky's strategy is to find a loosening of the Specified Subject Constraint that lets (88) in, but continues to block (83). He is guided by other apparent violations of the Specified Subject Constraint found in causative constructions.<sup>6</sup> Let's

<sup>6</sup> Chomsky (1992) cites the Government Transparency Corollary in Baker (1988) as his guide, one spe-

## 7. Subjects and Complex Predicates

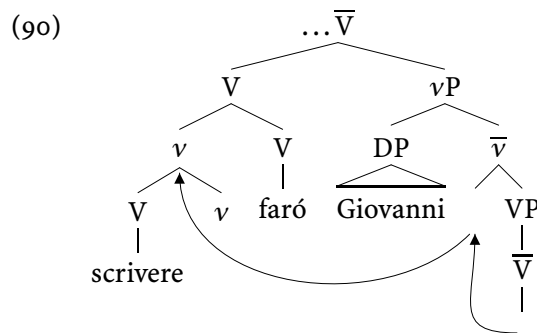
briefly consider these constructions and the way in which they are similar to the process we are confronted with in (83) and (88).

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.<sup>7</sup>

- (89) 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.'

This is striking because subjects are otherwise quite capable of surfacing pre-verbally 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 (90) below.<sup>8</sup> Indeed, not only may the subject of the lower clause not intervene between the



causative and lower verb, 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.

cial instance of which is found in the causative construction.

<sup>7</sup> My description of Italian causatives, and all the examples, comes from Burzio (1986).

<sup>8</sup> For an early proposal along these lines, see Aissen (1974).



- (91) 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 by terms in the higher clause. This is illustrated by the contrast in the following French examples (92).

- (92) a. Elle aurait fait tirer les pauvres soldats<sub>1</sub> l’un sur  
 she would have made shoot the poor soldiers the one on  
 l’autre<sub>1</sub>.  
 the other  
 ‘She would have made the soldiers shoot each other.’  
 b. \* Elles<sub>1</sub> auraient fait tirer le pauvre soldat l’une sur  
 they would have made shoot the poor soldier the one on  
 l’autre<sub>1</sub>.  
 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.<sup>9</sup>

So, on the basis of this evidence, let us adopt a picture of causatives which involves verb movement, as in (90). 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:

- (93) Alcuni prigionieri furono fatti lavorare.  
 a few prisoners were made work  
 ‘A few prisoners were made to work.’

<sup>9</sup> We will explore this description of the relationship between anaphors and their antecedents later.

## 7. Subjects and Complex Predicates

---

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  $\mu$  with accusative Case, this will mean that the subject of the lower VP has moved into the Specifier of  $\mu$ P in the higher clause to get Case. So, the tree in (90) should look instead something like (94) on the facing page. In this representation, the verb *scrivere* has adjoined to *faró*, and then they have together adjoined to  $\mu$  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  $\mu$ 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 (95).

- (95) Maria ha fatto riparare la macchina a Giovanni  
Mary has made repair the car to John  
'Mary has made John repair the car.'

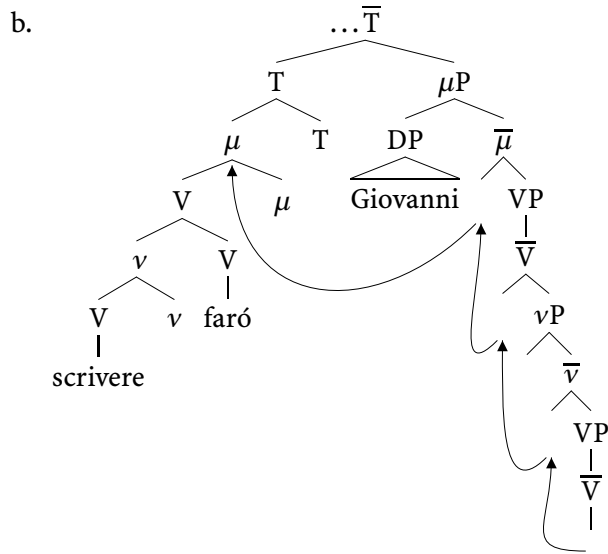
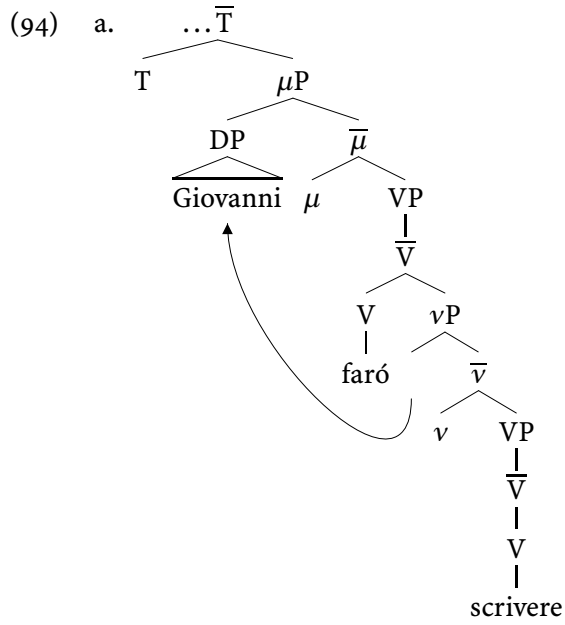
There are good reasons for thinking that the external  $\theta$ -role bearer in this example, *a Giovanni*, is in fact in a "subject" position. It blocks Cliticization of phrases that follow it, as (96) indicates, and this is just what the Specified Subject Constraint predicts if it is in subject position.

- (96) \* 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.'

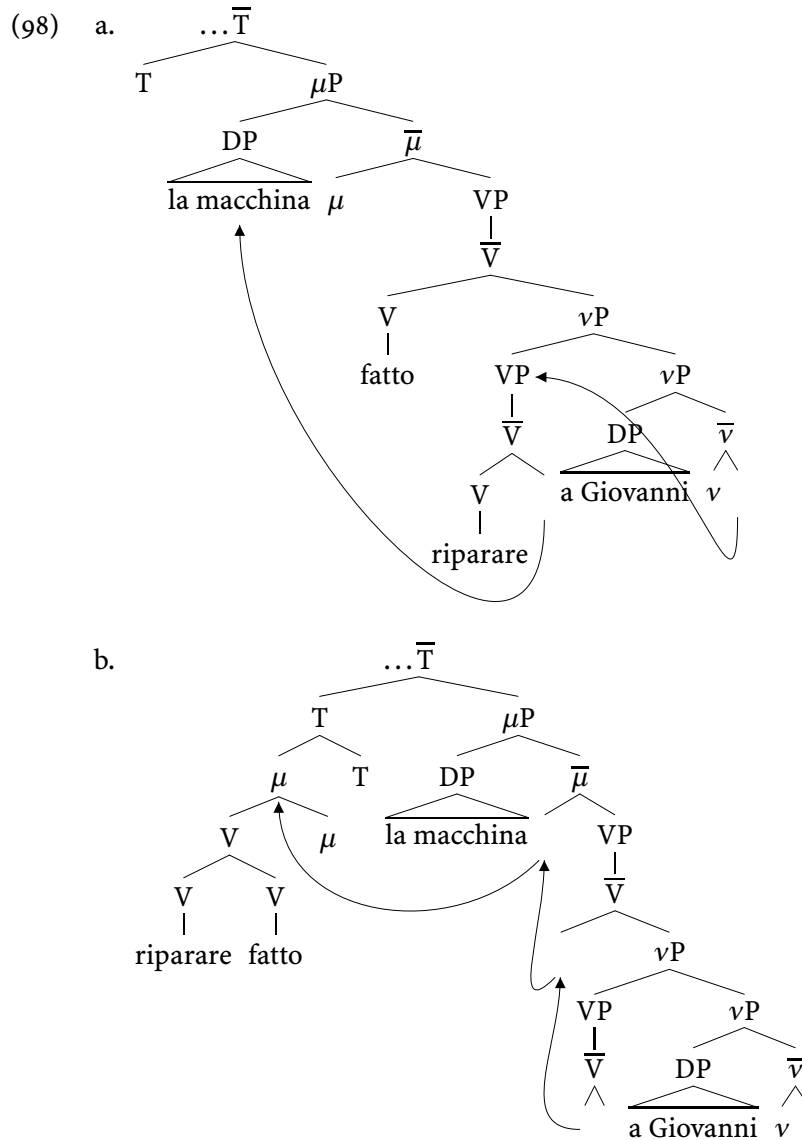
The reason for believing that the object, *la macchina*, is dependent on Case from something in the higher clause is that it's the one that loses accusative Case marking when the causative verb passivizes, as (97) indicates.

- (97) 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



## 7. Subjects and Complex Predicates



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 (95) a representation something like (98) above. 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 (96) cannot move into the higher clause.

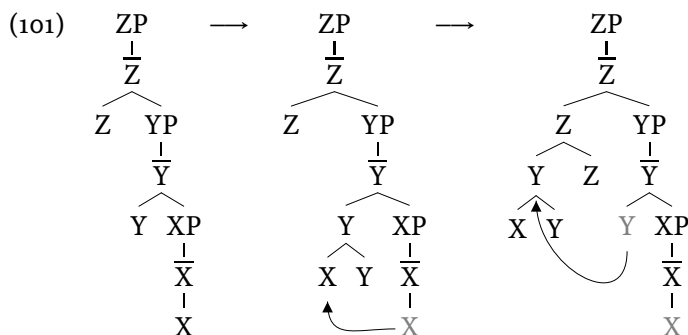
So perhaps instead we should parse these examples as in (99) on the following page, and seek a way of relaxing the Specified Subject Constraint so that accusative Case-marked complements are allowed to move past the subject, but other complements are not.

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  $\mu P$ . 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 (87b).

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. Furthermore, he adopts the view that movement steps create discontinuous versions of the things that are being moved. He calls these "chains." In the case of Head Movement, when one  $X^0$  adjoins to  $Y^0$ ,  $X^0$ 's chain is made up of the position that  $X^0$  moves into and the position it left. Let's represent this as  $(X^0, X^0)$ . Putting these two things together, we can derive (100).

(100)  $X^0$  Chains have at most two members.

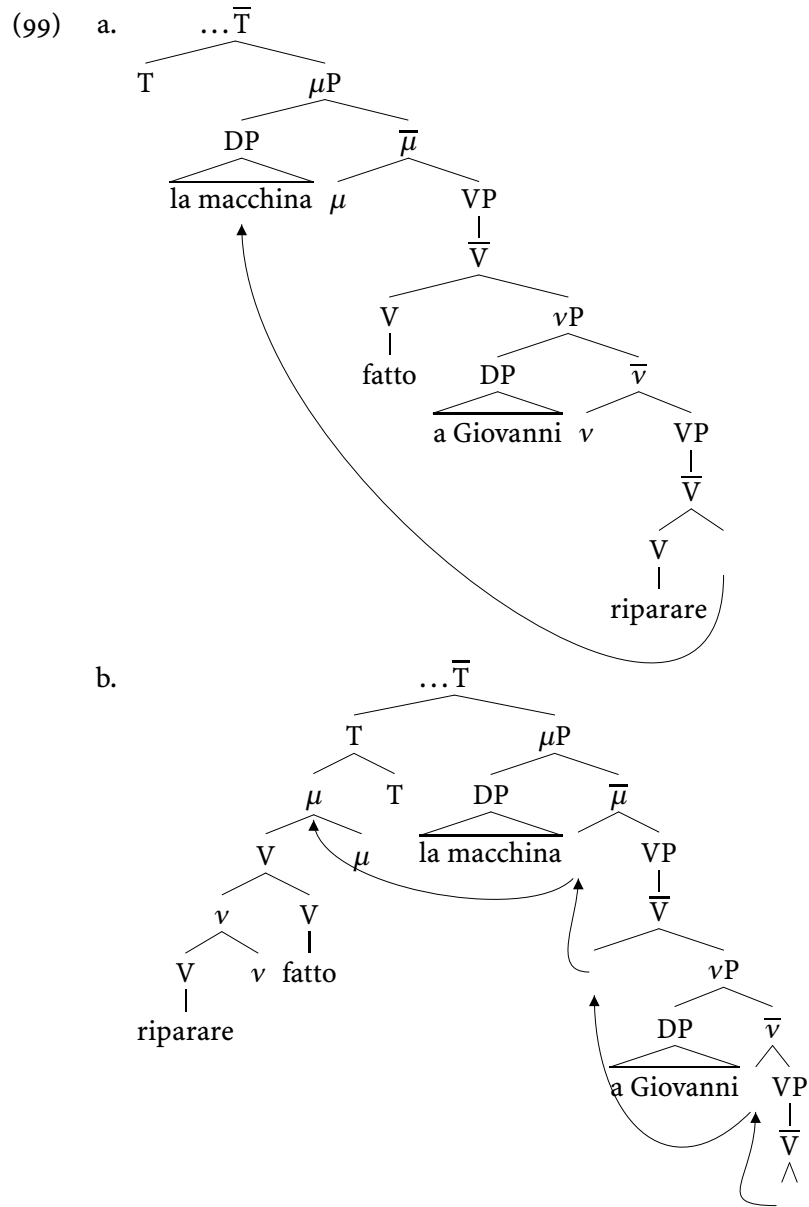
To see this, consider the derivation sketched in (101), in which a head,  $X^0$ , moves first to  $Y^0$  and then onto  $Z^0$ .



The last step in (101) 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

7. Subjects and Complex Predicates

---

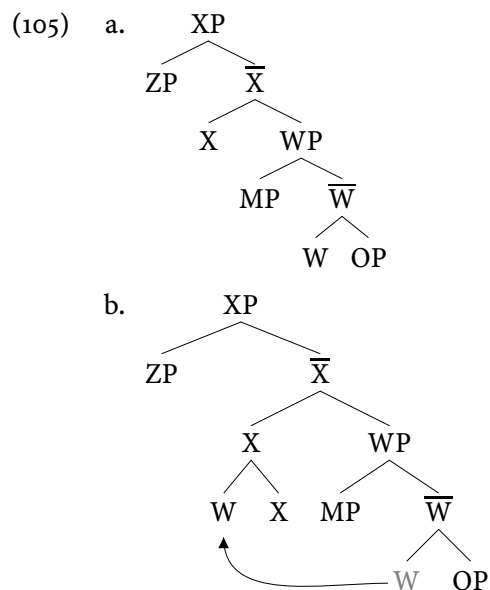


a different chain is created:  $(Y^0, Y^0)$ . In general, then, if Head Movement involves adjoining one head to another, then the derivation (101) represents the general case of iterative Head Movement. And, as can be seen, no chain will be formed that has more than two members.

Now the locality condition Chomsky formulates makes use of this idea about chains, and references a variety of terms which we define first.

- (102) Let the Domain of a chain ( $DOM(X, X)$ ) be the set of nodes contained within the smallest maximal projection containing the chain but not containing any  $\bar{X}$ .<sup>10</sup>
- (103) “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  $\alpha$ .<sup>11</sup>
- (104)  $\alpha$  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 there is no minimal S that contains  $\xi, \beta$ .

To see what these definitions do, consider the trees in (105).



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

<sup>10</sup> This is adapted from Chomsky 1992, p. 19.

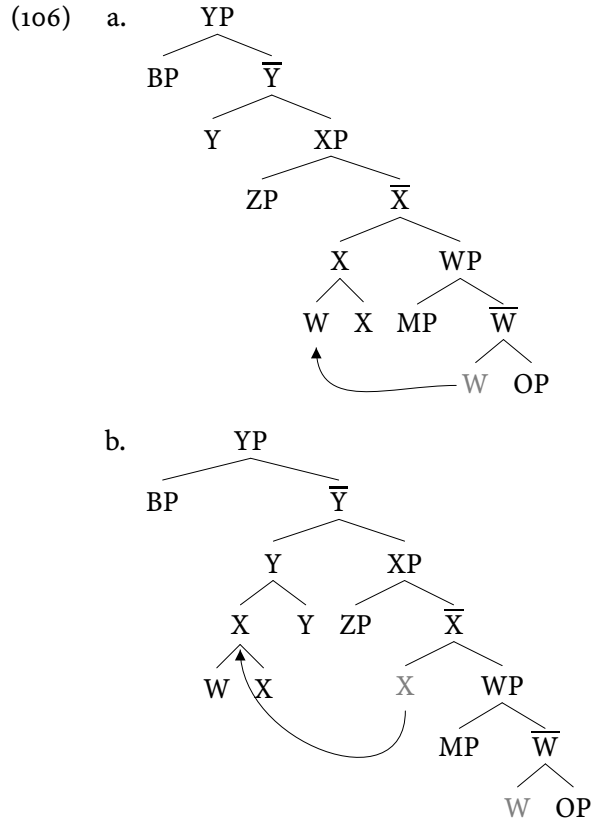
<sup>11</sup> Chomsky 1992, p. 16.

## 7. Subjects and Complex Predicates

{MP, OP}. The condition in (104) will now prevent OP from Argument 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 (105b). As before, the minimal domain of X is {ZP, WP}. But the Minimal Domain of (W,  $\bar{W}$ ) is going to be {ZP, MP, OP}; 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 it possible to Argument move OP to ZP.

Finally, consider a scenario in which there has been multiple head movement, as in (106).



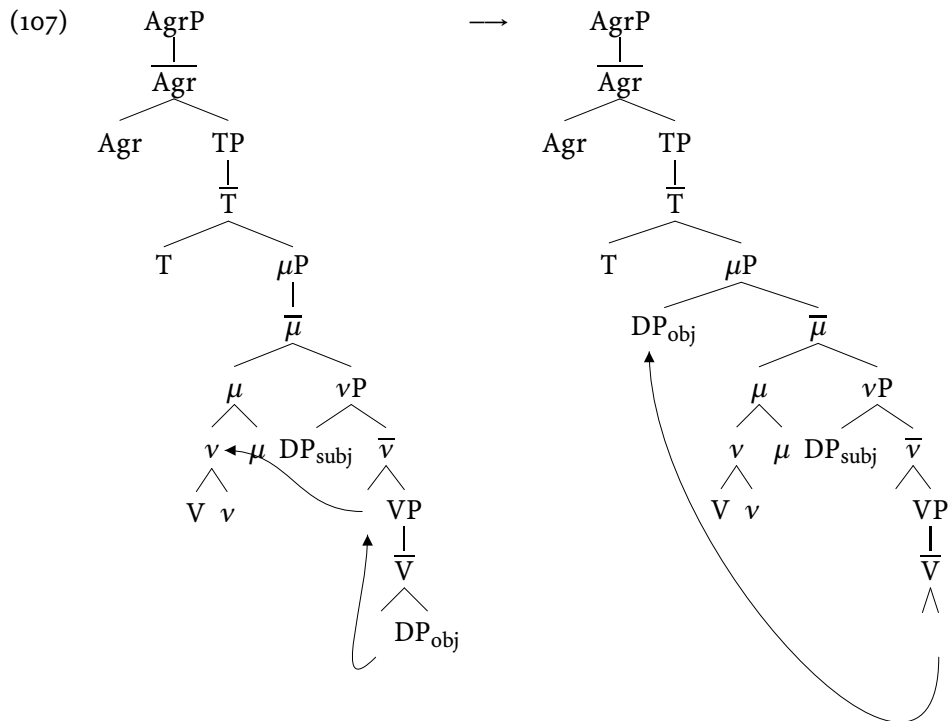
This derivation triggers in a crucial way the idea that chains which are formed by Head Movement have at most two members. It is this constraint that prevents (W,  $\bar{W}$ ) from having a larger minimal domain than that given to it in (105). Because the minimal domain of W doesn’t grow as a function of X’s movement, MP and BP will



not be in the same minimal domain, and as a consequence (104) will prevent OP from moving past both MP and ZP into BP.

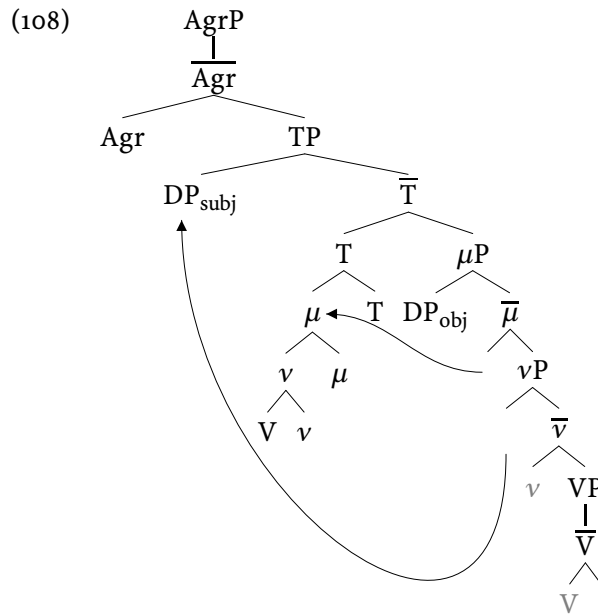
This is a general effect. And, as can be seen, it has the interesting consequence that it will derive the difference between the causative and non-causative clitic movement cases. Further, this proposal extends to our problem, and, interestingly, even might contribute to an explanation for why subjects 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. As (107) indicates, the portion of the derivation that brings the object past the subject will satisfy this locality condition.



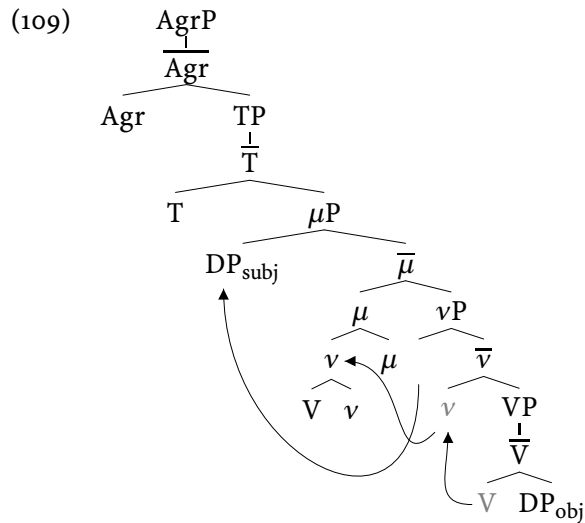
Because the minimal domain of  $(v, v)$  includes both the Specifier of  $vP$  and  $\mu P$ , movement of the object into Specifier of  $\mu P$  does not violate the locality constraint. The subsequent movement of the subject into Specifier of TP won't violate it either because adjoining  $\mu$  to T forms a chain whose minimal domain includes the Specifiers of TP and  $\mu P$ .

## 7. Subjects and Complex Predicates

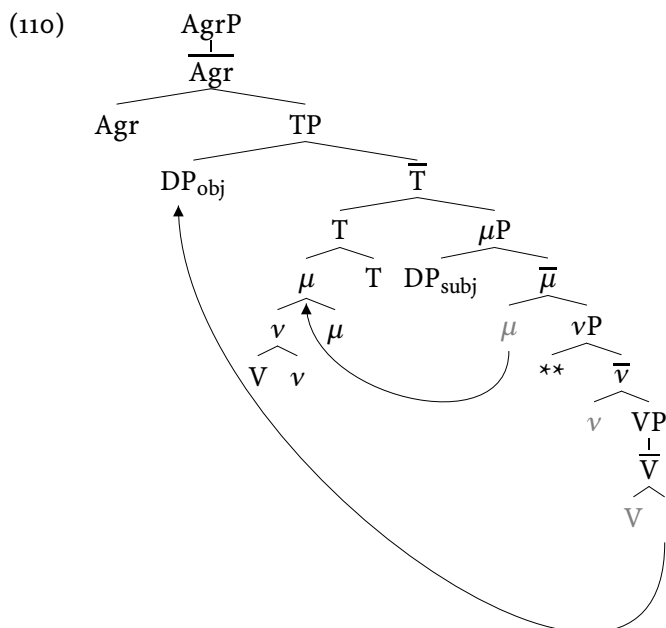


From the Specifier of TP, the subject now has a clear path to Specifier of AgrP, where it gets Nominative Case.

Consider by contrast a derivation that delivers the subject and object into the wrong positions. The subject will move into Specifier of  $\mu P$ , as indicated in (109).



Consider now the step that would bring the object into Specifier of TP.



It's this step that must be blocked. However once  $\mu$  adjoins to T, a minimal domain is created that includes Specifier of TP and Specifier of  $\mu$ P. For this reason, the movement of the object past  $DP_{subj}$  does not violate the minimal link condition. On the other hand, if there were something in Specifier of  $v$ P — the position marked with “\*\*” and from which the subject moved — then a violation would be created. This position is not in a minimal domain with Specifier of TP.

This account must be augmented with something that forces the position vacated by the subject to nonetheless be occupied by something. 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 actually occupies both positions. We shall soon see that there are other facts which seem to demand something like this idea.

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 might define movement, now, as in (111).

- (111) MOVE  $\alpha$
- a. Relocate  $\alpha$  from position  $\rho_1$  to position  $\rho_2$ .
  - b. Place in  $\rho_1$  a trace bound to  $\alpha$ .

## 7. Subjects and Complex Predicates

---

This has the nice auxiliary effect of offering a way to derive the Upwards Constraint. Bound variables 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. We will have an opportunity to examine this idea further in the coming chapters and see the reasons for believing it. Until then, however, let's provisionally accept this idea about Argument Movement.

The violation of the minimal link condition in (110) arises then because the object has moved past the trace of the subject. Since the position it moves into (Specifier of TP) is not in a minimal domain with the position the trace resides in (Specifier of  $vP$ ), this is prevented. In general, then, an object is blocked from moving past the subject if it skips Specifier of  $\mu P$ . The presence of a subject, then, forces an object to move into the position in which it gets Accusative Case.

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, however, as (112) indicates.

- (112) a. \* Him tried to speak.  
b. \* Her talked to Jimbo.  
c. \* Me ran.

Interestingly, however, this does seem to be possible in other languages. Unlike English, and all of IndoEuropean, which uses one Case for the subjects of transitive and intransitive clauses alike and has a different Case for objects, there are languages which use the same Case for objects and subjects of intransitive clauses, and a different Case for the subjects of transitive clauses. 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 be-

tween 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:

- (113) 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 (113). 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.

## 7.5 Remaining Problems

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 about English. Another is that it does not interact well with Sportiche's account of quantifier float. And finally, it seems to require a theory of Case assignment that puts the term marked with Case in the Specifier of the phrase that the Case assigner heads. While this is consistent with the genitive and nominative Case assignment, and we've now made it consistent with the accusative Case assigned to the direct objects of verbs, it is counterexemplified the accusative Case that is assigned to the subject position of infinitives like those in (114).

- (114) I prefer [<sub>CP</sub> for [ her to give the lecture]].

Let's examine each of these problems in turn.

## 7. Subjects and Complex Predicates

---

### 7.5.1 The main verb in English is too high

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 (115).

- (115) 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:

- (116) Jerry spoke fervently to Mary.

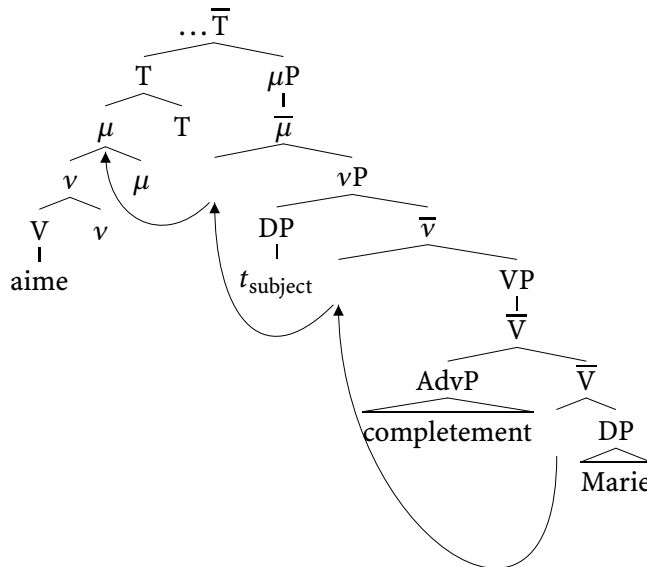
The contrast, therefore, targets something much narrower than the relationship between a main verb and its complement — it has to do with the relationship between a main verb and its accusative Case marked DP. What we look for, then, is something that forces the verb and Accusative DP to be adjacent 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:

- (117) a. Nothing can adjoin (to the left) of  $\mu P$ , and  
b. Specifier of  $\mu P$  is linearized to be first.

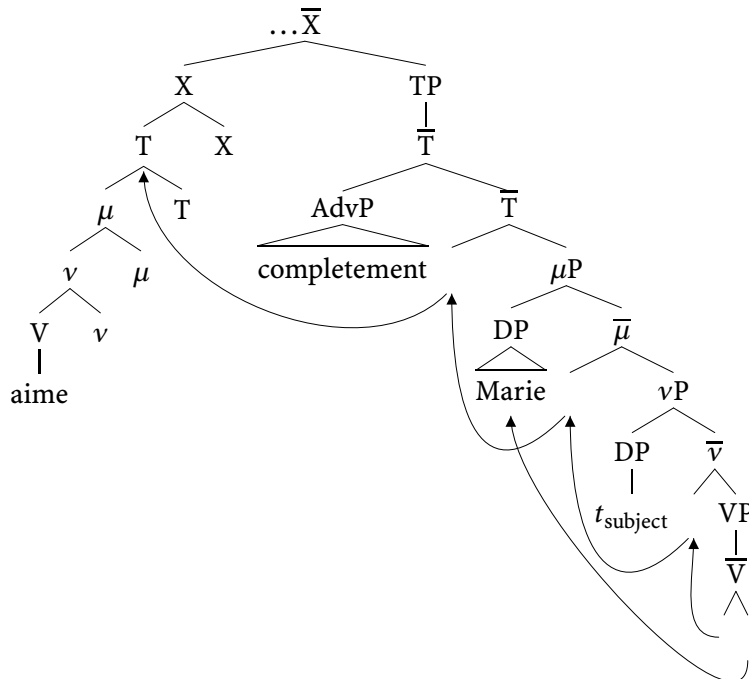
If this is the reason that Accusative DPs and verbs are adjacent, then there are many potential differences between English and French that could produce Pollock's difference. It could be, for instance, that French DPs are not forced to move overtly into Specifier of  $\mu P$  to get Case. This would 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 (115a) a parse like (118) below, for instance.

(118)



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 (116a) might have a representation like (119), in which  $X^0$  represents the new mystery head to which French verbs move.

(119)



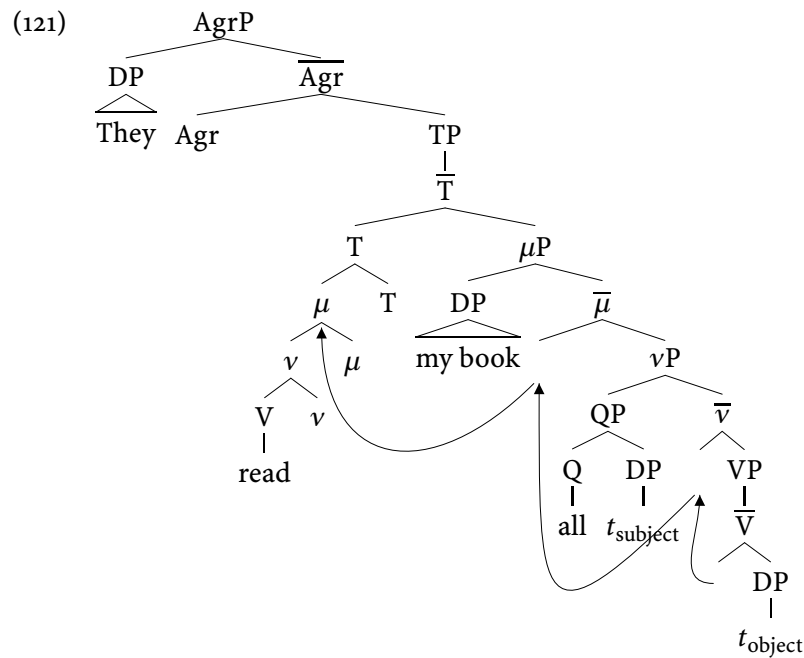
## 7. Subjects and Complex Predicates

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.

### 7.5.2 Incompatible with Quantifier Float

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 (120), whose parse is (121) below.

(120) \* 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 (122), and this too is incorrect.

(122) \*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.



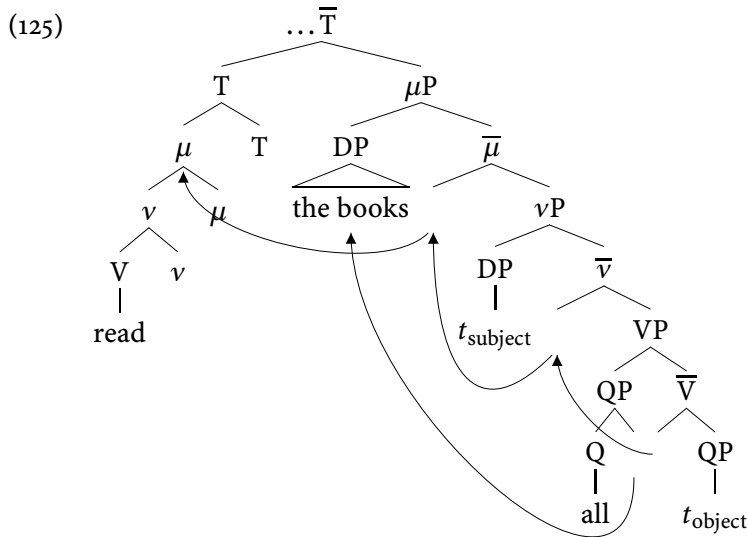
This difficulty is acknowledged in Sportiche's paper. He noted that his account predicts that objects which have moved in the Passive construction should be able to leave a floated quantifier in postverbal position, as in (123).

(123) \*The girls were visited all.

Sportiche's account requires something to block (123), and I suggest that it is also responsible for the ungrammaticality of (120) and (122). Something like (124) will get us a long way in this direction.<sup>12</sup>

(124) A quantifier cannot be floated in its c-selected position.

While this will be sufficient to prevent (120), it still leaves one derivation available to produce the ungrammatical (122) and (123). 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 (125) below. This derivation conflicts with

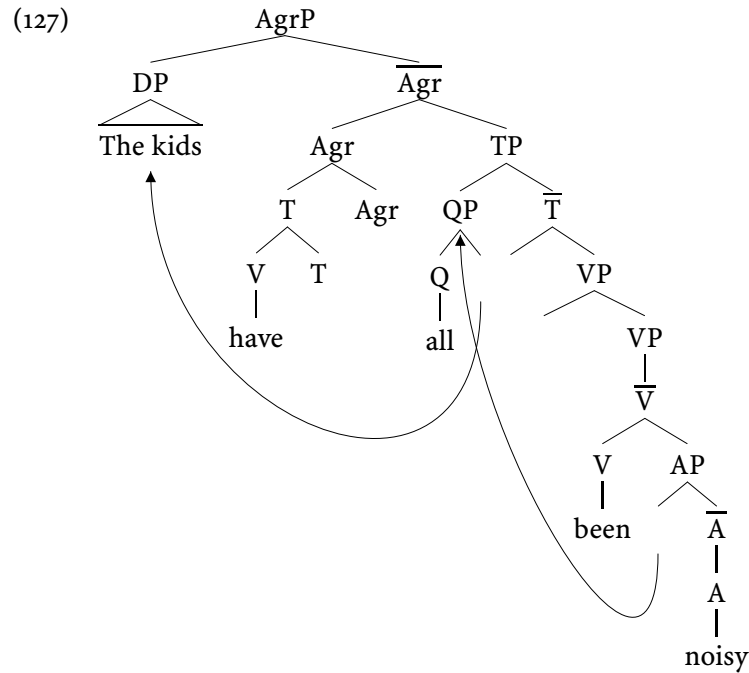


EARLINESS, however, as it is possible to create a shorter derivation that leads to a licit s-structure. Moving the object directly from its D-structure position to its Case marked, and therefore surface, position is a shorter derivation; EARLINESS will enforce that derivation over the one shown in (125). Unfortunately, as we've seen on the homework assignments, *Earliness* appears to be routinely violated in derivations leading to floated quantifiers. For instance, to produce the sentence in (126) requires the derivation sketched in (127).

<sup>12</sup> This is the conclusion reached in Déprez (1989) and Bošković (2001).

## 7. Subjects and Complex Predicates

(126) The kids have all been noisy.



EARLINESS should block this derivation, however, as there is a shorter one allowed in which the subject moves directly from its  $\theta$ -marked position into its surface position. If we are to use *Earliness* to block (125), then we should understand why it doesn't similarly block (126). Indeed, we should understand how (126) is produced if EARLINESS is valid no matter how we handle (125).

Here, then, are some problems we must leave hanging: what is the source of (124), and how are derivations like (125) to be blocked and (127) to be allowed.

### 7.5.3 PRO, Case Theory and the typology of Infinitives

The final problem 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 (128).

(128) 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>13</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 (128), that Case happens to be Accusative.

We initially adopted the view that it was *for* which assigns Case in (128) 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 “permeable” and “impermeable” infinitives involves, partly, the observation that only opaque infinitives are capable of housing PRO. We captured this by letting the distinction between “permeable” and “impermeable” infinitives devolve to whether they are CPs or AgrPs. Then we devised the condition on PRO in (129) that is sensitive to this distinction.


- (129) 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 (130) indicates.

- (130) 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 (128)/(130a).

In fact, there is something slightly odd about accounting for the distribution of PRO with (129). Not only is (129) 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 (131), PRO is driven from the object position of *promoted* into Specifier of infinitival AgrP.


- (131) Gary wanted [ PRO to be promoted ].  


<sup>13</sup> These ideas are a rough adaptation of Bošković (1996, 1997).

## 7. Subjects and Complex Predicates

---

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>14</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 (132).

(132) Gary wanted [ Sally to be promoted ] .  


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.

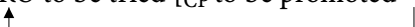
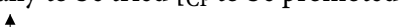
The way Argument Movement applies to PRO and overt DPs is really exactly the same. It is the same, in particular, in a way that suggests that we do not want a different engine driving them. Recall that we formulated Argument Movement as follows:

- (133) ARGUMENT MOVEMENT  
 Move an XP from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory, only if:
- i.  $\beta$  c-commands  $\alpha$ , and
  - ii. there is no CP that contains  $\alpha$  but not  $\beta$ , and
  - iii. the Specified Subject Constraint holds.

(147i) 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 (134), therefore.

- (134) ARGUMENT MOVEMENT  
 Move an XP from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory, only if:
- i. there is no CP that contains  $\alpha$  but not  $\beta$ ,
  - ii. and the Specified Subject Condition holds.

(134ii) captures the fact, recall, that Argument movement cannot take something out of a clause unless it's permeable. Thus, it captures the ungrammaticality of the cases in (135), where either a PRO or an overt DP has moved.

- (135) a. \* I wanted [ PRO to be tried [CP to be promoted ] ].  
  
 b. \* I wanted [ Sally to be tried [CP to be promoted ] ].  


<sup>14</sup> The EPP also requires an additional phrase marker to satisfy its requirements.

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 EARLINESS, which also applies to movement of PRO. EARLINESS suppresses movement of an argument unless doing so would not produce a speakable parse. In the case of overt DPs, for instance, it prevents movement from Case marked positions, as in (136).

(136) \*Jerzy seemed to            that Sal slept. (cf.: It seemed to Jerzy that Sal slept.)  
↑

(136) is ungrammatical because *Jerzy* moves from the position Case marked by *to*. Interestingly we need the same constraint on Argument Movement when it applies to PRO, as in (137) shows.

(137) \*I wanted [ PRO to seem to            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:

(138) \*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 (130c) are grammatical. In Icelandic, for instance, infinitival clauses with PRO subjects quite easily have an overt complementizer associated with them:

(139) *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. It is hard to see how to account for this contrast between English and Icelandic if the PRO Restriction determines the distribution of PRO. The placement

## 7. Subjects and Complex Predicates

---

of PRO in Icelandic otherwise appears to be the same as in English, so whatever fixes the placement of PRO in these languages should be minimally different. The PRO Restriction doesn't seem to allow for a weakening of the required sort.

If the PRO Restriction doesn't determine the distribution of PRO, then what does? There are two qualities that a replacement theory should have. It should explain why, roughly speaking, PRO is in complementary distribution with overt DPs. And it should also explain why movement of PRO seems to obey the same constraints that movement of overt DPs does. In particular, it should strive to explain the ungrammaticality of examples like (137).

The alternative we will explore has these two properties. It claims that PRO is subject to the Case filter in the same way that overt DPs are. PRO, however, bears a Case that no other DP can; let's call this "null Case," to be distinguished from the "overt Cases," Accusative, Nominative and Genitive.<sup>15</sup> The distribution of PRO, then, will be a function of the positions to which null Case is assigned. Thus, we should adopt a Case Theory like (140).

(140) CASE FILTER

- a. Every overt DP must occupy a "null" Case marked position and cannot occupy an "overt" Case marked position.
- b. Every PRO must occupy an "overt" Case marked position and cannot occupy a "null" Case marked position.

If we restrict the positions to which null Case is assigned to just the Specifiers of certain infinitival AgrPs, then we will capture why PRO is found only in these positions. This will have the desired effect of preventing PRO from occupying the Specifier position of finite AgrP, since this is a position to which an overt Case is assigned. In this way, (138) can be blocked. And this also decouples the position of PRO from the presence of complementizer, thus allowing situations like the Icelandic example in (139).

Thus, we have now at least four kinds of  $\text{Agr}^0$ :

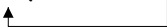
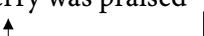
- (141)
- a.  $\text{Agr}_{\text{fin}}$  assigns Nominative.
  - b.  $\text{Agr}_{\text{inf}^a}$  assigns null.
  - c.  $\text{Agr}_{\text{inf}^b}$  assigns Accusative.
  - d.  $\text{Agr}_{\text{inf}^c}$  doesn't assign Case.

Verbs like *try* and *promise* select infinitival clauses headed by  $\text{Agr}_{\text{inf}^a}^0$ . Thus in these infinitivals we find PRO subjects, and never overt ones.

Consider now how this system will ensure that movement of PRO, like that of overt DPs, occurs just from non-Case marked positions. In the case of overt DPs



<sup>15</sup> The proposal that PRO is subject to the Case filter, and that the Case it bears is a proprietary one, is introduced in Chomsky and Lasnik, and explored systematically in Martin (2001).

this presently follows from Earliness. The movement in (142a) is blocked because the DP has moved farther than needed to satisfy the Case filter.

- (142) a. \* Jerry seems to that chocolate is good.  

  
 b. Jerry was praised .  


By contrast, in the example of passive in (142b) movement of the overt DP is permitted because it is required by the Case filter. Earliness prevents the needlessly long derivation in (142a) but allows the derivation in (142b) precisely because it is required.

These derivations are distinguished in a slightly different way when movement of PRO is involved. The passive example in (143b) is allowed by Earliness for the same reason that (142b) is.

- (143) a. \* I want PRO to seem to that chocolate is good.  

  
 b. I want PRO to be praised .  


This derivation is necessary to satisfy the Case filter's requirement on PRO. But unlike (142a), (143a) is also permitted by Earliness. Because PRO does not get null Case in its underlying position, movement to Specifier of AgrP is made necessary by the Case filter and therefore allowed by Earliness. What blocks (143a) is the stipulation that PRO cannot be assigned an overt Case. In its D-structure position PRO is assigned the overt Accusative Case, and this is responsible for violating this stipulation. In the derivation in (143a), by contrast, PRO is assigned only one Case — its underlying position in this circumstance is not assigned any Case.

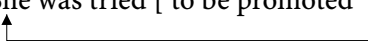
An important role is played, then, in the stipulation that PRO is incapable of receiving overt Case. This is a stipulation that goes beyond merely making PRO susceptible to the Case filter. And, as we shall see, there is a role played by the parallel stipulation in (140a) that overt DPs are similarly prevented from receiving null Case. The thesis here then is not only that PRO, like overt DPs, is subject to the Case filter but also that there are two unmixable Cases: overt and null. We should seek an explanation for why these Cases cannot be mixed. We might think of null Case as being, in fact, just like overt Case, but with the phonological consequence that the DP that bears it be silent. Overt Case, by contrast, we should interpret as being a Case whose phonological reflex is pronouncability. An overt DP cannot bear null Case because it is not silent, and PRO cannot bear overt Case because it is silent.

## 7. Subjects and Complex Predicates

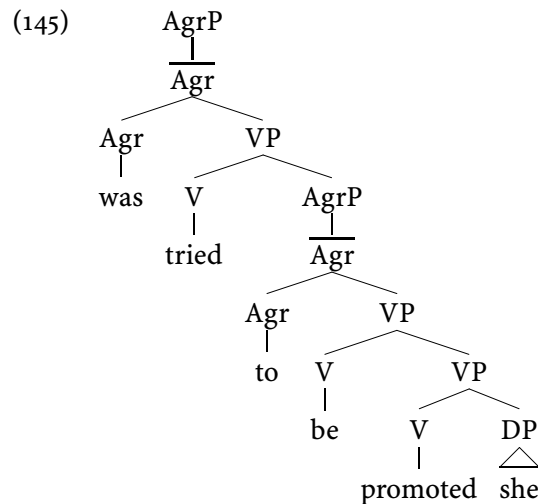
---

One consequence of this proposal is that 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 CPs does now is to prevent Argument Movement from relocating things out of them. But this can now be derived from (140) and the Extended Projection Principle. To see this, consider how (144), an example of the putative constraint against A Movement out of CPs, will be derived.

(144) \* She was tried [ to be promoted ].



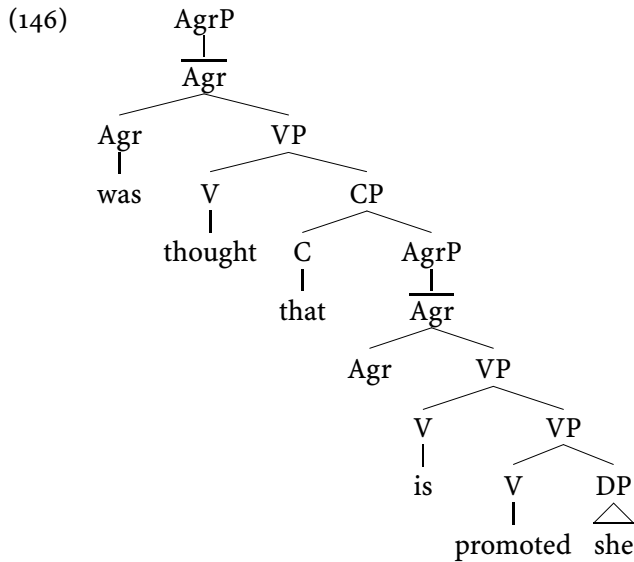
If we let the infinitival complement be just an AgrP, and not a CP, then we can take the underlying representation to be as in (145). (I've foreshortened lots of steps here — in particular, much of the structure belonging to the root clause has been suppressed.)



The Extension to the Projection Principle is going to require that both AgrPs have something in their Specifier position at some point in the derivation. The only thing that can satisfy that requirement is *she*. This means that *she* will have to move first into one of these Specifiers and then into the other. But both these Specifiers are Case marked, now: the lower one gets null Case from the infinitival Agr<sup>0</sup>, and the higher one gets Nominative Case from the finite Agr<sup>0</sup>. But moving *she* into the Specifier of the lowering AgrP will give it null Case, and this violates the requirement that overt DPs receive only overt Case. Thus this case will be prevented in a way that is parallel to how (143a) is blocked.

Movement out of finite CPs will also be blocked by a combination of the EPP and the conditions on Case. Consider, by way of illustration, a D-structure like that in (146) on the facing page. The EPP will again require a derivation from this D-





structure that brings *she* into both Specifier of AgrPs. But when *she* moves into the lower Specifier of AgrP it will satisfy the Case filter, and Earliness will prevent it from moving from this position.

In general terms, then, argument movement from AgrPs whose Specifiers are assigned Case will be blocked. We no longer need to stipulate that CPs cannot be escaped by Argument Movement. In particular, impermeable infinitives can be defined 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 Agr<sup>0</sup>s can assign is the null Case, we capture the fact that PRO is capable of living in only impermeable clauses. Thus we capture one of the correlations that the permeable/impermeable distinction was defined over: impermeable infinitives can host PRO but are island for Argument Movement, where as permeable infinitives cannot host PRO and are not islands for Movement.

So, we can simplify the definition of Argument Movement once more, to:

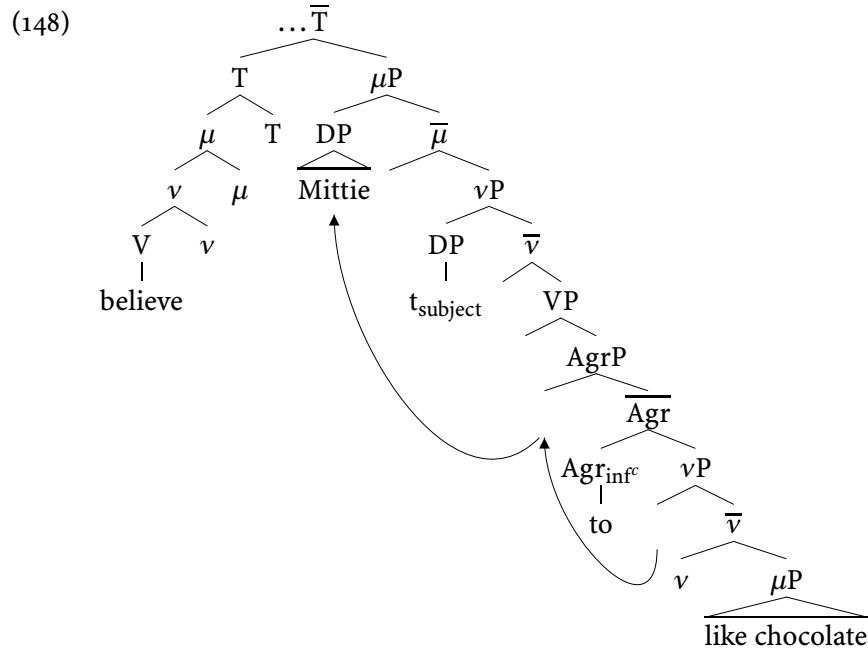
(147) ARGUMENT MOVEMENT

Move an XP from  $\alpha$  to  $\beta$ ,  $\beta$  an empty position licensed by  $\bar{X}$  Theory.

What about permeable infinitives? These come in two sorts, neither of which allow PRO inside them. These, then, are infinitives headed by Agr<sup>0</sup><sub>inf<sup>c</sup></sub>: the Agr<sup>0</sup> 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

## 7. Subjects and Complex Predicates

infinitives, in other words, will not be islands for Argument movement. When these infinitives are c-selected by a verb that assigns an external  $\theta$ -role, then there will be an Accusative Case assigning  $\mu$ 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 (148) below.

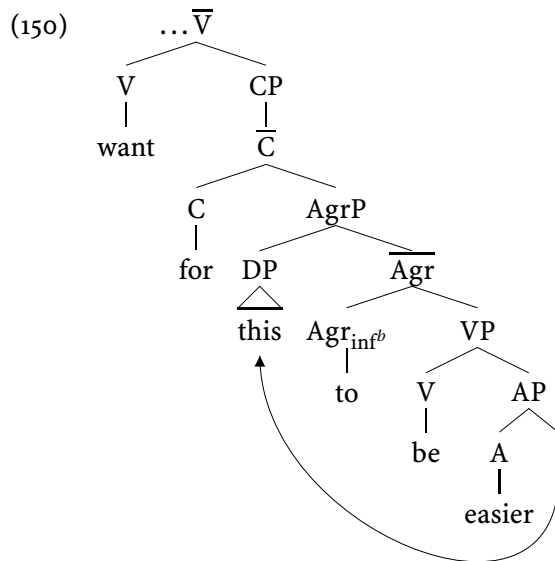


The other sort of verb that c-selects a permeable infinitive is *seem*, a verb that has neither an external  $\theta$ -role nor  $\mu$ P. As a consequence, in these cases, if there is an argument within the infinitive Case it will not be able to move into the  $\mu$ P immediately above the infinitival clause, as happens in (148), 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.

Permeable 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.

- (149)
- a. Permeable infinitives =<sub>def</sub> AgrPs headed by an Agr<sup>0</sup> that does not assign Case.
  - b. Impermeable infinitives =<sub>def</sub> AgrPs headed by an Agr<sup>0</sup> that assigns Case.

What's left is to characterize the kinds of infinitives that are selected by verbs such as *want* and *prefer*. These are the infinitives that can show up with the complementizer *for*, and which allow Accusative subjects. This is where we encounter the final kind of Agr<sup>0</sup>, the one that assigns Accusative Case. This AgrP is, apparently, only selected by the complementizer *for* in English, and it is a particular kind of impermeable infinitive, then. If we let Agr<sup>0</sup><sub>inf<sup>b</sup></sub> represent the Accusative Case assigning Agr<sup>0</sup>, then this case will get a representation like (150) below. Because the Specifier



of this infinitive is marked with Accusative Case, the Case based account of PRO's distribution will prevent PRO from being in this position. In this way, then, is the ungrammaticality of (151) derived.

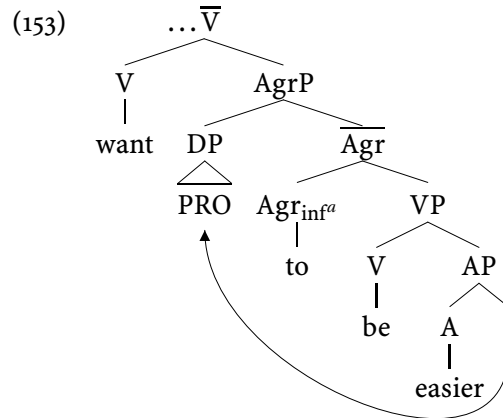
(151) \*I wanted for PRO to be easier.

What about the other forms that this infinitive can take? Recall that it can host PRO if the complementizer is absent, as in (152a), or surface with an overt Accusative DP in this situation, as in (152b).

- (152) a. I want [PRO to be easier].  
 b. I want [this to be easier].

I suggest that we treat (152a) as a normal opaque infinitive, and (152b) as a normal permeable 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 (152a) will get is parallel to that given to the complements of *try* or *promise*; it's something like (153) (where Agr<sup>0</sup><sub>inf<sup>a</sup></sub> represents the Agr<sup>0</sup> that assigns null Case).

## 7. Subjects and Complex Predicates



(152b) will get a representation parallel to that we've given to *believe* in (148). This is different than the original account we entertained for this case. We adopted the view that the infinitive in (152b) is an impermeable 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:

- (154) a. Mittie was believed to like chocolate.  
 b. \* This was wanted to be easier.

If (152b) really is a permeable infinitive, then this difference must be accounted for in some other way.

The advantage to seeing (152b) as a permeable 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*.

- (155) 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, as (156) indicates.

- (156) 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 (155c) and (156c) by way of this difference between adjectives and verbs if the

Accusative Case that is assigned in (155c) comes from the higher clause. This is just what classing these as permeable 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 null Case. These predicates, *promise*, *try*, *glad*,<sup>16</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>16</sup> As, for instance, in “I am glad to meet you.”



# 8

## Rethinking Things

We have a lot of balls in the air, at the moment. It will be useful, perhaps, to pause 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. Let's begin by bringing together the central elements of the grammar.

### 8.1 Review

That portion of the grammar formerly expressed with phrase structure rules is now expressed by way of the following set of well-formedness conditions.

- (1)  $\bar{X}$  Skeleton:
  - a.  $XP \rightarrow \{\alpha, \bar{X}\}$
  - b.  $\bar{X} \rightarrow \{\bar{X}, \beta\}$
  - c.  $\bar{X} \rightarrow \{X^0, \gamma\}$
  
- (2) Linearization Parameters:
  - a. Specifier: [first, last]
  - b. Projection of  $X^0$ : [first, last], modulo (2a)

## 8. Rethinking Things

---

- (3)  $[\alpha, \beta] =_{\text{def.}} \alpha$  precedes  $\beta$ .  
 $\{\delta \alpha, \beta\} =_{\text{def.}} \overbrace{\delta}^{\alpha} \beta$  or  $\overbrace{\delta}^{\beta} \alpha$ .
- a. For all words,  $x$  and  $y$ , within a phrase marker, either  $[x, y]$  or  $[y, x]$ .  
 b. 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  $Y$ .
- (4) **PROJECTION PRINCIPLE**  
 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  $\alpha$ .
- (5) **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.
- (6) **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 and Theta Criterion conflict with the Case filter.

- (7) **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.

- (8) **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  $\overline{X}$  Skeleton, the Projection Principle, the Theta Criterion and the Modification rule.
  - c. Let *s-structure* be a phrase marker that satisfies the Case Filter and is phonologically interpreted.



- d. A *Derivation* =<sub>def</sub> an ordered n-tuple of phrase markers,  $\mathcal{D} = (P_1, P_2, \dots, P_n)$ , such that:
- i. Each  $P_i = \mathcal{R}(P_{i-1})$ , and
  - ii. The first member is a D-structure, and
  - iii. There is at least one member that is an s-structure, and
  - iv. There is some  $P_i$  in which all inflectional features are checked.
  - v. For every AgrP,  $\alpha$ , in the D-structure, there is some phrase marker which has a phrase in Specifier of  $\alpha$ .<sup>1</sup>
- (9) A grammatical sentence must have a well-formed derivation.
- (10) EARLINESS  
Let  $\mathcal{D} = \{D_1, D_2, \dots, D_n\}$  be the set of well-formed derivations for some sentence, S, and  $\mathcal{O}$  be the set of d's, such that for every  $D_i \in \mathcal{D}$ ,  $d_i$  is that sub-series of parses in  $D_i$  that starts with the D-structure of  $D_i$  and ends with the first s-structure in  $D_i$ . The s-structure of S is the one in the shortest  $d \in \mathcal{O}$ .
- (11) Instances of  $\mathcal{R}$
- i. A Movement: Move an XP.  
Subject to the “minimal distance” version of the SSC.
  - ii. Head Movement: Move an  $X^0$ .  
Subject to the Head Movement Constraint.

Every  $\mathcal{R}$  is subject to:

- (12) 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 (1995b) 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 (15).

<sup>1</sup> This is the Extension to the Projection Principle.

## 8. Rethinking Things

---

- (13) Let every DP have a strong Case feature, and let  $\mu$  and Agr have Case features as well.
- (14) 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.
- (15) 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, (13) 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, (14) 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 (16).

- (16) Let *s-structure* be a phrase marker that has no strong features and is phonologically interpreted.

Chomsky (1995b) has suggested tying the phonologically interpreted part together with the strong feature part more directly with:

- (17) A phrase marker with strong features cannot be phonologically interpreted.

And we change (8div) so that it reads:

- (18) Some  $P_i$  is phonologically interpreted.

We might similarly think of expressing the EPP, i.e. (8dv), 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 (8dv) with:

<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.

(19) 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 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:

(20) 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, the Theta Criterion, the Likes-Attracts-Likes Constraint and Earliness. We dispense with the Case Filter, putting in its place an feature-based interpretation of Case. 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. Our definition of a derivation and transformational rules now looks like this:

(21) 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  $D$ -structure be a phrase marker with lexical items that satisfies the  $\bar{X}$  Skeleton, Projection Principle, 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. Some  $P_i$  is phonologically interpreted.

<sup>3</sup> The "real" Relativized Minimality inspires this, and can be found in Rizzi (1990).

## 8. Rethinking Things

---

- (22) Instances of  $\mathcal{R}$ :
- i. A Movement: Move XP into a position licensed by  $\bar{X}$  Theory.
  - ii. Head Movement: Move  $X^0$  and adjoin it.

### 8.2 Towards Deriving $\bar{X}$ Theory

#### 8.2.1 Kayne's "Antisymmetry" hypothesis

There are certain generalizations about the derivations we've examined so far. One of these is that all movement 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 must be in the scope of the phrase that binds them to 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:

- (23) 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 (23). And (23) 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,

<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.

a smattering of western African languages (e.g., Eβe, Abe), and no doubt many others.

Kayne (1994) seeks to derive this trend. What he proposes is a linearization algorithm that will replace the linearization parameters (2) by strengthening (3) so that it universally fixes the position of Specifiers, Complements and Adjuncts.

Presently (3), repeated here, has the role of fixing the linear order of terminals once the linearization parameters have set the linear order of phrases.

- (3)  $[\alpha, \beta] =_{\text{def.}} \alpha \text{ precedes } \beta.$   
 $\{\delta \alpha, \beta\} =_{\text{def.}} \begin{array}{c} \delta \\ \alpha \quad \beta \end{array} \text{ or } \begin{array}{c} \delta \\ \beta \quad \alpha \end{array} .$
- a. For all terminals,  $x$  and  $y$ , within a phrase marker, either  $[x, y]$  or  $[y, x]$ .
  - b. 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  $Y$ .

The linearization parameters will impose a language-particular ordering on Specifiers, Adjuncts and complements which (3) will map on a linearization of the terminals those phrases contain. Thus, for instance, the string in (24) is given the structure in (25) by the  $\bar{X}$  Skeleton and other constraints on  $D$ -structure.<sup>5</sup>

(24) the coin under the chair

(25) {the, {{coin}, {under, {{{the, {chair}}}}}}}}

The linearization parameter settings for English will impose on this structure the ordering in (26).

(26) [ the, [ coin, [ under, [ the, chair ]]]]

What this says is that *the* precedes the phrase *coin under the chair*, and that *coin* precedes the phrase *under the chair*, and that *under* precedes the phrase *the chair*, and finally that *the* precedes the phrase containing *chair*. Nothing in (26) fixes the relative order of *the* and *coin*, nor the relative order of *coin* and *chair*, and so on. This is what (3) does. It says that if  $\alpha$  precedes  $\beta$  then all of the terminals in  $\alpha$  precede all the terminals in  $\beta$ . Thus from (26) it derives the complete linearization in (27).

(27) [the, coin], [the, under], [the, the], [the, chair], [coin, under], [coin, the], [coin, chair], [under, the], [under, chair], [the, chair]

<sup>5</sup> I leave out the detail of NumP, and the other functional projections within a DP, in this example.

## 8. Rethinking Things

---

This is how our present theory works.

Note that the linearization algorithm in (3), like all successful linearization algorithms, has the consequences in (28).

- (28)
- 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).

The algorithm in (3) has these consequences, at any rate, if phrase markers are binary branching (i.e., a tree that has no more than two daughters per node), because the Linearization parameters will linearize every sister, and (3) will linearize all the terminals for each sister.

Our present theory makes the linearization of terminals in a phrase marker a slave to the linearization parameters. The only thing (3) adds, beyond meeting the requirements in (28), 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.

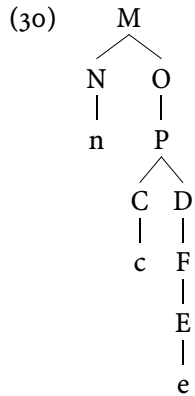
Kayne's idea is to reverse this. He proposes to make the way that sisters, and consequently terminals, are linearized a function of asymmetric  $c$ -command. First, he interprets the statements in (28) as constraints on a linearization,<sup>6</sup> and then defines a linearization as follows.

- (29)
- 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.  $d(X) =_{\text{def}}$  the set of all terminals dominated by  $X$ .
  - d.  $d(\langle X, Y \rangle) =_{\text{def}}$  the set of all ordered pairs  $\langle x, y \rangle$  such that  $x$  is dominated by  $X$  and  $y$  is dominated by  $Y$ .
  - e.  $d(\langle X_1, Y_1 \rangle, \langle X_2, Y_2 \rangle, \dots, \langle X_n, Y_n \rangle) =_{\text{def}}$   $d(\langle X_1, Y_1 \rangle) \cup d(\langle X_2, Y_2 \rangle) \dots \cup d(\langle X_n, Y_n \rangle)$ .
  - 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.

<sup>6</sup> Rather than, say, a consequence of the linearization procedure.

He calls (29g) the *Linear Correspondence Axiom* (LCA).

Let's see this work in an example:



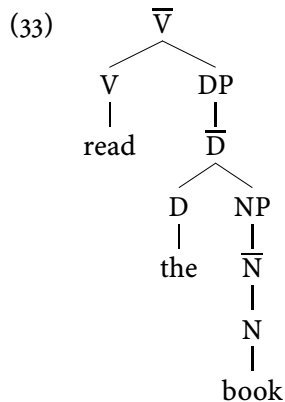
For this phrase marker we find the following values for A and d(A):

- (31) 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.  $d(A) = \{ \langle n, c \rangle, \langle n, e \rangle, \langle c, e \rangle \}$

(31b) 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, (31) produces either (32a) or (32b).

- (32) a.  $\{ [n, c], [n, e], [c, e] \}$   
 b.  $\{ [c, n], [e, n], [e, c] \}$

A structure like that in (30) which might be found in nature is:



Which the LCA will linearize in one of the ways in (34).

## 8. Rethinking Things

---

- (34) 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 (35).

- (35) a. 
$$\begin{array}{c} C \\ \diagup \quad \diagdown \\ D \quad E \\ | \quad | \\ d \quad e \end{array}$$
 b. 
$$\begin{array}{c} C \\ \diagup \quad \diagdown \\ D \quad E \\ | \quad | \\ F \quad G \\ | \quad | \\ f \quad g \end{array}$$

In both of these cases, the linearizations fail. As (36) shows, (35a) is (very) not total, and (35b) is not antisymmetric.

- (36) a. (35a):  
 $A = \{ \}$   
 $d(A) = \{ \}$
- b. (35b):  
 $A = \{ \langle D, G \rangle, \langle E, F \rangle \}$   
 $d(A) = \{ \langle f, g \rangle, \langle g, f \rangle \}$

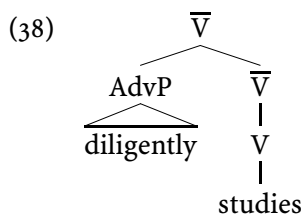
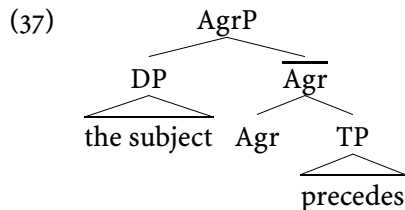
Banning (35a) is welcome when C is not an  $X^0$ . This would prevent two heads from being sisters — that is, it will prevent a phrase from having two heads. Where we might want to allow (35a) is when one  $X^0$  has adjoined to another. Actually, we haven't seen many cases of this in which two terminals are involved — but we might want to consider instances of cliticization to involve such a structure. Let's come back to this.

Banning (35b) 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 (35b) 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 in each case 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:

(39)  $\alpha$  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 (39), but his suggestion targets the case in (35a), when C is a head, as well. In order to unify these two situations, he proposes that the relation a Specifier has to its sister phrase is the same as that which an adjunct has to its sister. He proposes, in other words, that there is no distinction between XPs and  $\bar{X}$ s. This decision will undermine some of our principles (the modification rule, for instance), but let's be brave and forge ahead nonetheless. So, in place of (1), we have:

- (40) a.  $XP \rightarrow \{XP, YP\}$   
 b.  $XP \rightarrow \{X^0, (YP)\}$

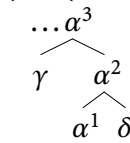
And, over the structures formed by these, we define these terms:<sup>7</sup>

<sup>7</sup> The notions "category" and "segment" come from May (1985), and were used extensively in Chomsky (1986a).

## 8. Rethinking Things

---

(41) A CATEGORY is the set of  $\alpha$ 's in:  $\alpha^n$  such that each  $\alpha^i$  is a projection



of  $\alpha^{i-1}$ , and  $\alpha^1$  is  $\alpha^0$ . Each  $\alpha^{i>1}$  is a SEGMENT of that category.

Kayne's method of allowing the phrase markers in (35) uses this segment/category distinction by way of making c-command sensitive to it, as in (42).

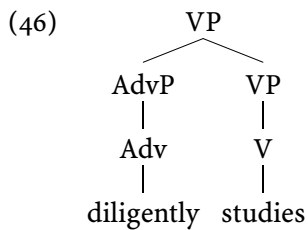
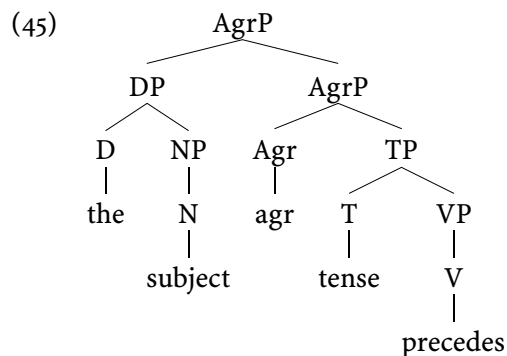
(42)  $\alpha$  c-commands  $\beta$  iff  $\alpha$  and  $\beta$  are categories and every category that dominates  $\alpha$  dominates  $\beta$ , and  $\alpha$  excludes  $\beta$ .

(Kayne 1994, (3), p. 16)

(43) A category  $\alpha$  EXCLUDES  $\beta$  iff no segment of  $\alpha$  dominates  $\beta$ .

(44)  $\alpha$  DOMINATES  $\beta$  if every segment of  $\alpha$  contains  $\beta$ .

Let's consider how this will work in the problematic situations in (37) and (38), which would now have the representations in (45) and (46).



When we compute out the A and d(A) of these phrase markers, we find:

(47) For (45):  
 $A = \{ \langle D, N \rangle, \langle DP, AgrP \rangle, \langle DP, Agr \rangle, \langle DP, TP \rangle, \langle DP, T \rangle, \langle DP, VP \rangle, \langle DP, V \rangle, \langle Agr, T \rangle, \langle Agr, VP \rangle, \langle Agr, V \rangle, \langle T, V \rangle \}$

$$d(A) = \{ \langle the, subject \rangle, \langle the, agr \rangle, \langle the, tense \rangle, \langle the, precedes \rangle, \langle subject, agr \rangle, \langle subject, tense \rangle, \langle subject, precedes \rangle, \langle agr, tense \rangle, \langle agr, precedes \rangle, \langle tense, precedes \rangle \}$$

- (48) For (46):  
 $A = \{ \langle AdvP, V \rangle, \langle AdvP, VP \rangle \}$   
 $d(A) = \{ \langle diligently, studies \rangle \}$

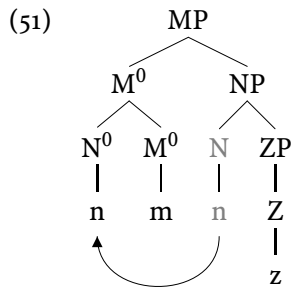
These are total, transitive and antisymmetric linearizations, and will yield one of the outcomes given in (49) below.

- (49) For (47):  
 $\{ [the, subject], [the, precedes], [subject, precedes] \}$   
 $\{ [subject, the], [precedes, the], [precedes, subject] \}$

- (50) For (48):  
 $\{ [diligently, studies] \}$   
 $\{ [studies, diligently] \}$

Notice that in (45), 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 (46) 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.

The same effects described here for phrases carry over to heads. If one head adjoins to another, as in (51), then they will be linearized in an antisymmetric way.



The A and d(A) for (51) are as in (52).

- (52)  $A = \{ \langle N, M \rangle, \langle N, N \rangle, \langle N, ZP \rangle, \langle N, Z \rangle, \langle M, N \rangle, \langle M, ZP \rangle, \langle M, Z \rangle, \langle N, Z \rangle \}$   
 $d(A) = \{ \langle n, m \rangle, \langle n, n \rangle, \langle n, z \rangle, \langle m, n \rangle, \langle m, z \rangle \}$

## 8. Rethinking Things

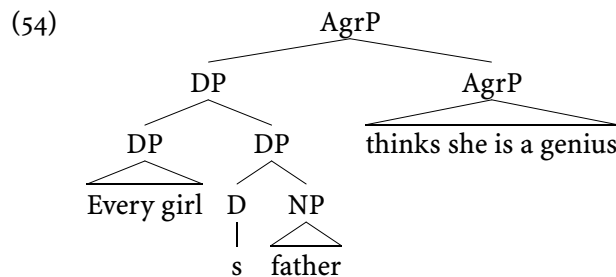
---

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 (42) 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 (53).

(53) 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 (42), and the  $\bar{X}$ -free phrase markers we are examining, this indeed obtains, as (54) below indicates.



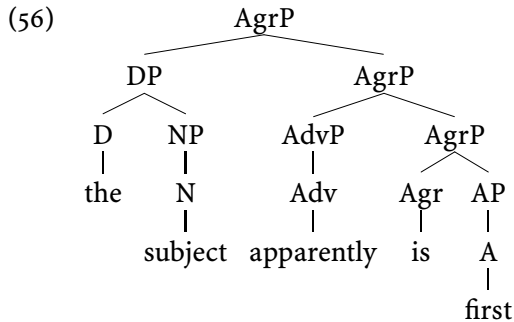
But there are other data that suggest that this is not a consequence we desire for c-command. Neither of the other two diagnostics of scope we have relied on suggest that the genitive DP is capable of c-commanding out of the DP it is embedded within. There is no disjoint reference effect in (55a), nor can the anaphor be bound in (55b).

- (55) 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 (42) 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 (56).<sup>8</sup>

<sup>8</sup> As in many of the phrase markers in this chapter, I've ruthlessly truncated structure needless to the point at hand.



The A and d(A) of (56) are given in (57) below. This violates antisymmetry, by

- (57)  $A = \{ \langle D, N \rangle, \langle DP, AgrP \rangle, \langle DP, Adv \rangle, \langle DP, Agr \rangle, \langle DP, AP \rangle, \langle DP, A \rangle, \langle AdvP, D \rangle, \langle AdvP, NP \rangle, \langle AdvP, N \rangle, \langle AdvP, AgrP \rangle, \langle AdvP, Agr \rangle, \langle AdvP, AP \rangle, \langle AdvP, A \rangle, \langle Agr, A \rangle \}$

$d(A) = \{ \langle the, subject \rangle, \langle the, apparently \rangle, \langle subject, apparently \rangle, \langle apparently, the \rangle, \langle apparently, subject \rangle, \langle apparently, is \rangle, \langle apparently, first \rangle, \langle is, first \rangle \}$

virtue of  $\langle the, apparently \rangle$  and  $\langle apparently, the \rangle$ , among others. In essence, then, this way 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 (58).

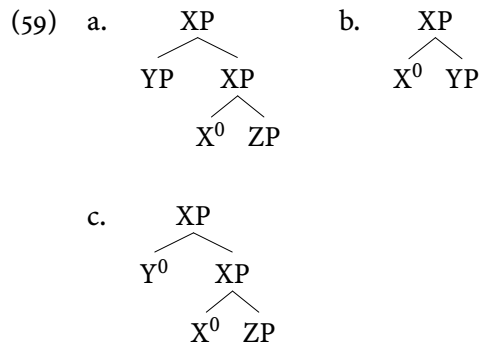
- (58)  $\alpha P \rightarrow \{ \alpha', (\beta) \}$ , where  $\alpha'$  is  $X^0$  or XP.

This, together with the LCA and the definition of category and segment, will allow just the shapes in (59) on the following page. The first two of these are well attested, of course; but (59c) needs to be blocked. Kayne suggests a principle which insists that Specifiers be phrasal. In any case, if some way of blocking (59c) 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 start? How does the LCA derive the fact that “up” translates as “precedes” in so many cases? Kayne suggests deriving this from

## 8. Rethinking Things

---



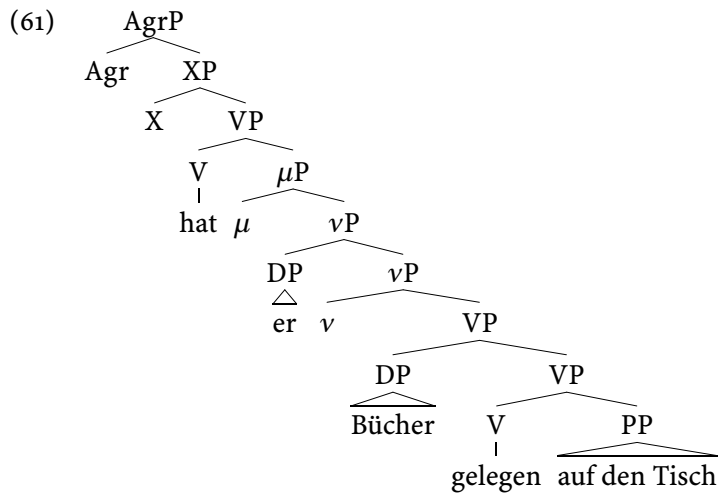
the LCA and the stipulation that  $d(A)$  always translates into a “precedes” linearization; that is:  $\langle \alpha, \beta \rangle =_{\text{def.}} [\alpha, \beta]$ . Once the force that blocks (59c) is identified, only (59a,b) are permitted and if  $\langle \alpha, \beta \rangle =_{\text{def.}} [\alpha, \beta]$  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.

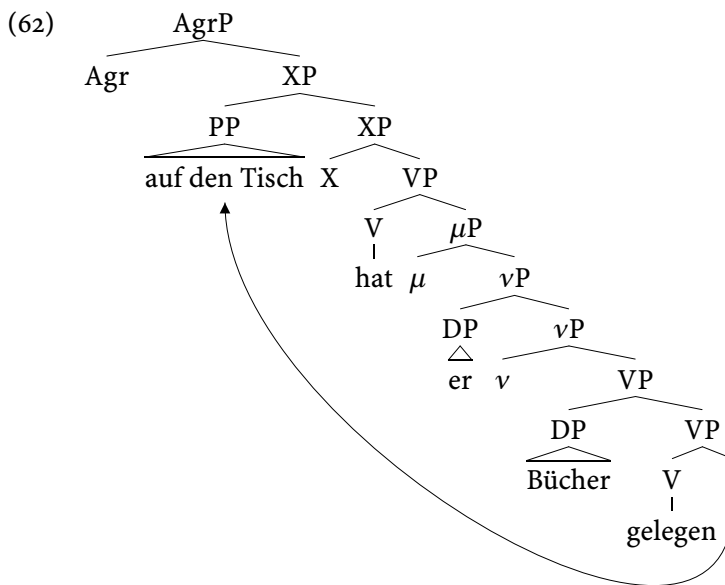
Of course, Kayne is aware of German, and the many other languages which appear to be head final. He suggests that these languages are, indeed, head initial underlyingly, but that a suite of movement operations systematically obscures this at s-structure. To get an image of how this might work, let’s reconsider German word-order from this perspective. We might give the sentence whose surface form is (60) an underlying representation like that in (61).<sup>9</sup>

- (60) ... weil Er auf den Tisch Bücher gelegen hat  
 ...since he on the table books put has  
 ‘...since he has put books on the table’

304 <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.



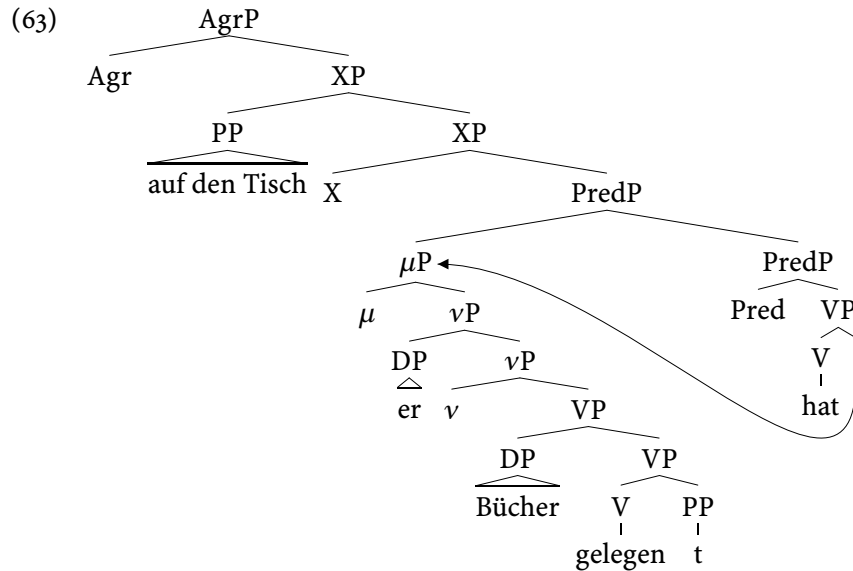
Imagine that Scrambling is obligatory, and relocates *auf den Tisch* to XP, as indicated in (62).



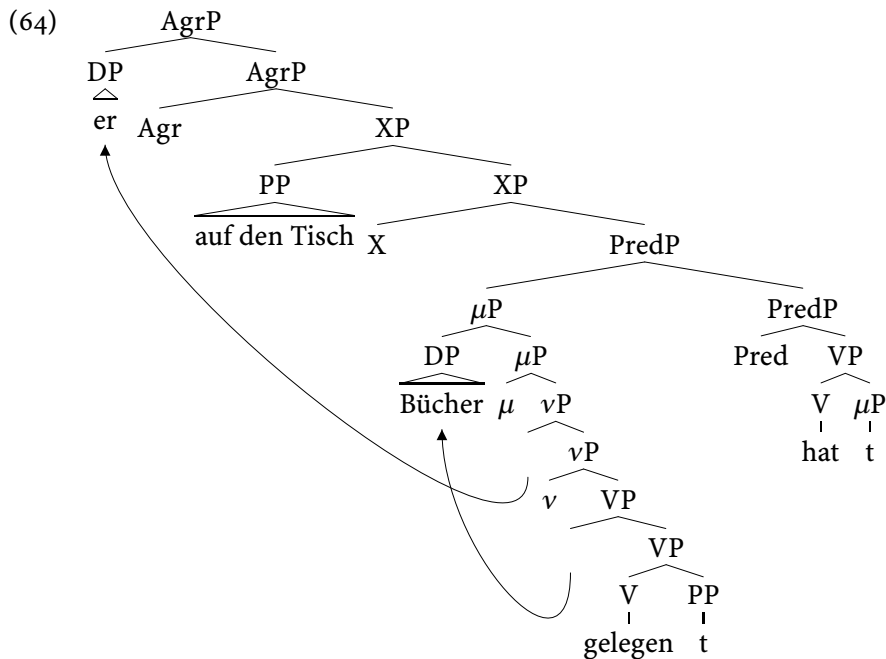
Making Scrambling obligatory, then, has the potential of bringing all the complements to the left of the verb that *c*-selects them. Now what's necessary is to get  $\mu P$  to precede *hat*. This can be achieved by moving  $\mu P$  leftwards. Zwart (1997) argues for something along these lines, and suggests that there is a functional phrase —

## 8. Rethinking Things

“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 (63).



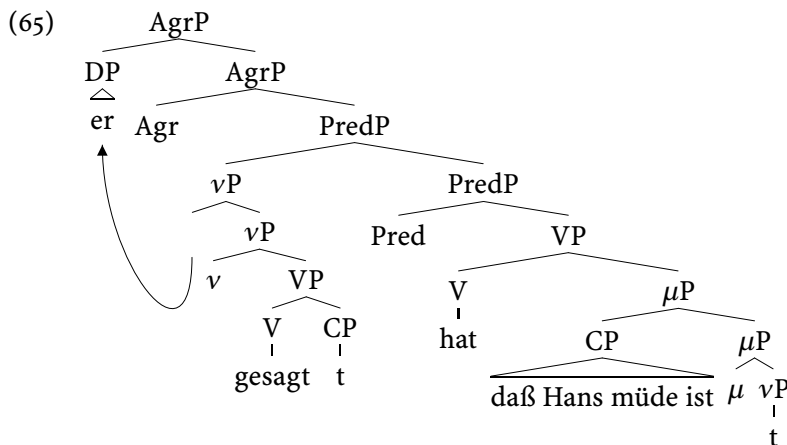
Once the two DPs have moved into their Case marked positions, we'll get the correct word order, shown in (64).



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 (65).

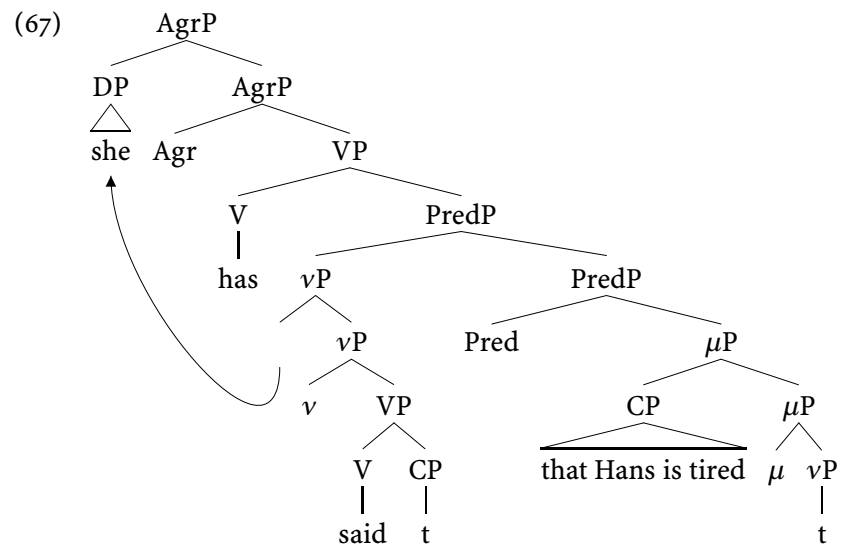
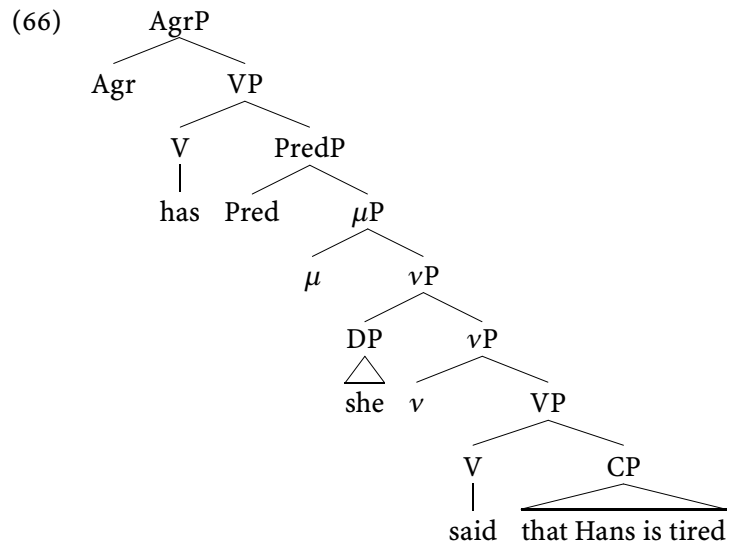


These derivations may look odd, but they would have to be mirrored in English as well in contexts where NP Shift arises. Because Kayne rigs the LCA so that it prevents rightward movement, the phenomena that “NP Shift” describes cannot be characterized in the terms that Ross did. On the other hand, Larson’s 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 (66) on the next page. 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 (65) will create a form like (67) on the following page in English. 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 force 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

8. Rethinking Things

---



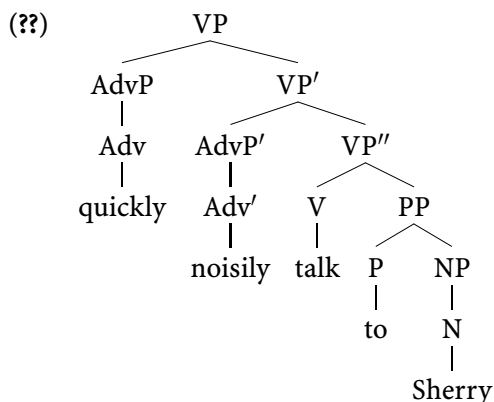
some puzzles which find an explanation on this view. To date, much less success has been garnered in rigidly head final languages, such as Korean and Japanese, however. This remains an open project.

Kayne's proposal for capturing the correlation between precedence and c-command, then, requires 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 (35). (That is, for the fact that on a standard definition of c-command, the LCA prevents any 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:

- (68)  $\alpha$  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 (??), 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

## 8. Rethinking Things

result, the linearization that results is not antisymmetric, and fails. But under (68) 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 (??) are as in (69), and satisfy the complete, transitive and antisymmetric conditions on a linearization.

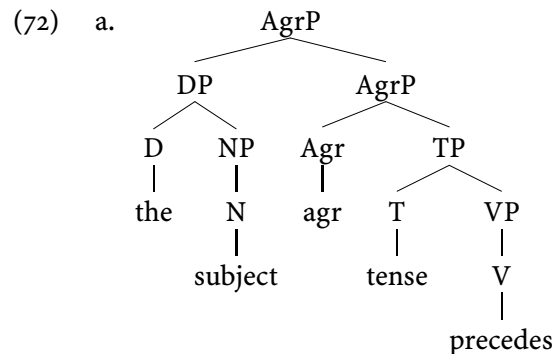
- (69) a.  $A = \{ \langle \text{AdvP}, \text{VP}' \rangle, \langle \text{AdvP}, \text{AdvP}' \rangle, \langle \text{AdvP}, \text{Adv}' \rangle, \langle \text{AdvP}, \text{VP}'' \rangle, \langle \text{AdvP}, \text{V} \rangle, \langle \text{AdvP}, \text{PP} \rangle, \langle \text{AdvP}, \text{P} \rangle, \langle \text{AdvP}, \text{NP} \rangle, \langle \text{AdvP}, \text{N} \rangle, \langle \text{AdvP}', \text{VP}'' \rangle, \langle \text{AdvP}', \text{V} \rangle, \langle \text{AdvP}', \text{PP} \rangle, \langle \text{AdvP}', \text{P} \rangle, \langle \text{AdvP}', \text{NP} \rangle, \langle \text{AdvP}', \text{N} \rangle, \langle \text{V}, \text{P} \rangle, \langle \text{V}, \text{NP} \rangle, \langle \text{V}, \text{N} \rangle \}$
- b.  $d(A) = \{ \langle \textit{quickly}, \textit{noisily} \rangle, \langle \textit{quickly}, \textit{talk} \rangle, \langle \textit{quickly}, \textit{to} \rangle, \langle \textit{quickly}, \textit{Sherry} \rangle, \langle \textit{noisily}, \textit{talk} \rangle, \langle \textit{noisily}, \textit{to} \rangle, \langle \textit{noisily}, \textit{Sherry} \rangle, \langle \textit{talk}, \textit{to} \rangle, \langle \textit{talk}, \textit{Sherry} \rangle, \langle \textit{to}, \textit{Sherry} \rangle \}$

We might also consider the possibility that the problem for the LCA which (35) poses could trace back to something other than the conditions the LCA imposes. Perhaps the linearization algorithm simply can't see the phrases in these examples which create the problem. If we pursued this direction, it wouldn't be necessary to meddle with  $c$ -command to account for why the sisters to Specifiers and adjoined heads do not enter into the LCA's computation. Imagine that we revert to something very close<sup>10</sup> to our original definition of  $c$ -command (in (70)), and then posit (71).

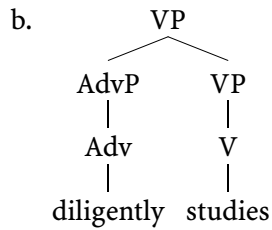
- (70)  $\alpha$   $c$ -commands  $\beta$  iff every phrase that dominates  $\alpha$  dominates  $\beta$ ,  $\alpha$  excludes  $\beta$  and  $\beta$  doesn't dominate  $\alpha$ .

- (71) Only categories enter into the linearization algorithm.

Consider how the problematic cases work out under this scenario. In cases where a phrase is in Specifier position, or an adjunct is adjoined to a phrase, such as (72), the  $A$  and  $d(A)$  will be as in (73).



<sup>10</sup> This is empirically equivalent to our original definition, but adds that  $\beta$  doesn't dominate  $\alpha$ . This is done to remove linearization paradoxes that arise if a phrase could  $c$ -command something that dominates it. So far as I can see it doesn't have any untoward outcomes with respect to characterizing scope phenomena.



(73) a. For (72a):

$A = \{ \langle DP, AgrP \rangle, \langle DP, Agr \rangle, \langle DP, TP \rangle, \langle DP, VP \rangle, \langle DP, T \rangle, \langle DP, V \rangle, \langle D, N \rangle, \langle Agr, T \rangle, \langle Agr, VP \rangle, \langle Agr, V \rangle, \langle T, V \rangle \}$

For (72b):

$A = \{ \langle AdvP, VP \rangle, \langle AdvP, V \rangle \}$

b. For (72a):

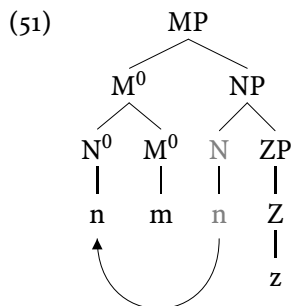
$d(A) = \{ \langle the, subject \rangle, \langle the, agr \rangle, \langle the, tense \rangle, \langle the, precedes \rangle, \langle subject, agr \rangle, \langle subject, tense \rangle, \langle subject, precedes \rangle, \langle agr, tense \rangle, \langle agr, precedes \rangle, \langle tense, precedes \rangle \}$

For (72b):

$d(A) = \{ \langle diligently, studies \rangle \}$

These are the desired linearizations, and they satisfy the constraints of completeness, transitivity and antisymmetry.

Consider next how this set of assumptions would apply to cases where one  $X^0$  has adjoined to another, as in (51), repeated below.



To evaluate the linearization algorithm for this case requires first recognizing that c-command is defined in terms of dominating phrases, and not dominating  $X^0$ s. As a consequence, when one head is embedded within another, then it will c-command

## 8. Rethinking Things

---

all the terms within the smallest phrase that contains both of them. With this consideration in mind, the A and d(A) that (70) and (71) manufacture for (51) are (74).

- (74) a.  $A = \{ \langle N, M \rangle, \langle N, N \rangle, \langle N, ZP \rangle, \langle N, Z \rangle, \langle M, ZP \rangle, \langle M, Z \rangle \}$   
 b.  $d(A) = \{ \langle n, m \rangle, \langle n, n \rangle, \langle n, z \rangle, \langle m, n \rangle, \langle m, z \rangle \}$

The linearization in (74b) is complete, transitive and antisymmetric. Note that N c-commands M, but that M doesn't c-command N, since it doesn't exclude it. For this reason N asymmetrically c-commands M, and *n* is linearized relative to *m*. In general, then, this set of definitions would force heads adjoined to another head to be linearized so that they precede the head they adjoin to.

This scheme would also allow for recursive adjunction of phrases. For instance, in (??), where two adverb phrases are adjoined recursively to VP, the A and d(A) would be as in (75).

- (75) a.  $A = \{ \langle AdvP, AdvP' \rangle, \langle AdvP, Adv' \rangle, \langle AdvP, V \rangle, \langle AdvP, PP \rangle, \langle AdvP, P \rangle, \langle AdvP, NP \rangle, \langle AdvP, N \rangle, \langle AdvP', V \rangle, \langle AdvP', PP \rangle, \langle AdvP', P \rangle, \langle AdvP', NP \rangle, \langle AdvP', N \rangle, \langle V, P \rangle, \langle V, NP \rangle, \langle V, N \rangle, \langle P, N \rangle \}$   
 b.  $A = \{ \langle \textit{quickly}, \textit{noisily} \rangle, \langle \textit{quickly}, \textit{talk} \rangle, \langle \textit{noisily}, \textit{to} \rangle, \langle \textit{noisily}, \textit{Sherry} \rangle, \langle \textit{noisily}, \textit{talk} \rangle, \langle \textit{talk}, \textit{to} \rangle, \langle \textit{talk}, \textit{Sherry} \rangle, \langle \textit{to}, \textit{Sherry} \rangle \}$

This linearization meets the requirements of completeness, transitivity and anti-symmetry, and, moreover, correctly linearizes the terminals.

Although this modification of Kayne's program is unsatisfying in leaving the stipulation in (71) unexplained, I suggest that we adopt it from now on. It has the important virtue of leaving the definition of c-command that is known to work for scope effects undamaged. And it does so while preserving not only Kayne's account of why higher tends to equate with precedes but also his method of deriving some of the constraints imposed by the  $\bar{X}$  Skeleton.

It also preserves the result that Specifiers and heads must be linearized in the same way with respect to their sisters. If, as Kayne suggests, " $\langle \alpha, \beta \rangle$ " is always interpreted as making  $\alpha$  precede  $\beta$ , then this will allow only head initial and Specifier initial phrases. Note, however, that there is nothing in the LCA that requires this interpretation. Kayne's proposals allow us to manipulate how " $\langle \alpha, \beta \rangle$ " is interpreted until the facts we desire emerge. That Specifiers precede their sisters, and the concomitant fact that phrasal movement is leftwards, does seem to be a cross-linguistic trend worth trying to derive. The head-initial order of phrases, and the concomitant fact that head movement is always leftwards, by contrast, does not transparently have the same cross-linguistic force. We might imagine controlling how " $\langle \alpha, \beta \rangle$ " is interpreted to reflect this. For instance, we might imagine a scheme like that in (76) which would permit right headed languages that have initial Specifiers.

- (76) For  $\alpha$  a head:  
 $\langle \alpha, \beta \rangle =_{\text{def.}} [\beta, \alpha]$  (Korean/Japanese)  
 $\langle \alpha, \beta \rangle =_{\text{def.}} [\alpha, \beta]$  (English/German?)

For  $\alpha, \beta$  phrases:  
 $\langle \alpha, \beta \rangle =_{\text{def.}} [\alpha, \beta]$  (Universal?)

If this is correct, we might strive to understand why the linearization of heads has a freedom that the linearization of phrases does not.

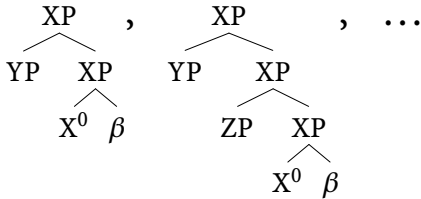
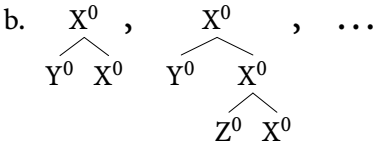
The theory Kayne (1994) provides us, then, has some flexibility. It is possible to preserve his account for the interesting trend that asymmetric c-command seems to map onto precedence, and even derive the simplification of the  $\bar{X}$  Skeleton, without embracing all of the consequences for phrase markers that he suggests.

### 8.2.2 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 (77).

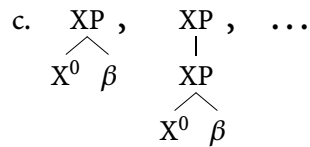
- (77)  $\alpha P \rightarrow \{\alpha', (\beta)\}$ , where  $\alpha'$  is  $\alpha^0$  or  $\alpha P$ .

The LCA, along with some decisions about how to define c-command, 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 system in (70) and (71), rather than Kayne’s, the LCA and (77) permit only the shapes in (78), where “ $\beta$ ” is, in each case, optional.

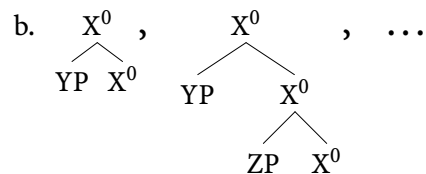
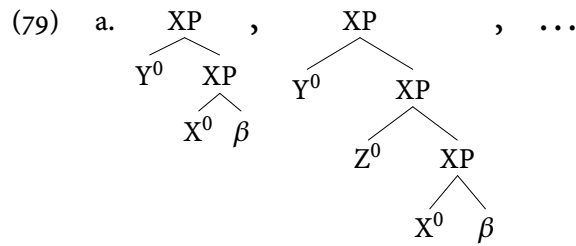
- (78) a. 
- b. 

## 8. Rethinking Things

---



If we let (77) have the role that the  $\bar{X}$  Skeleton had in defining D-structures, then D-structures will be confined to the forms in (78). Although Kayne let the LCA hold of every phrase marker in a derivation, if (77) 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 (79), and, at least so far as we've seen up to now, these should be prevented.



Since these structures would be formed by a movement operation adjoining a head to a phrase (in (79a)) or a phrase to a head (in (79b)), 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 (77) 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.<sup>11</sup> 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 ideas.

- (80)
- a. The notion of *projection*, e.g., an  $X^0$  projects an XP.
  - b. The *segment/category* distinction (see (41)).
  - c. A distinction between heads ( $X^0$ ) and phrases (XP).
  - d. The notion of *sets*, or *dominates*, that (77) expresses, i.e., that one, or two, terms can form another.

The heart of Chomsky’s idea is to express the distinction in (80c) in terms of (80d). This, in turn, will require a new formulation of *projection*.

He begins by following Speas (2000) 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 to think of  $\bar{X}$  theory as providing a template into which lexical insertion sticks terminals.

As an alternative, 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 (80d) makes reference to. Moreover, following the idea in Muysken (1982) that inspired also Kayne, 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 (77), then, Chomsky proposes a set forming operation that is defined over terminals. He calls this operation MERGE.

- (81)  $\text{MERGE}(x)(y) =_{\text{def.}} \{x, y\}$ , where  $x$  and  $y$  are terminals or  $\{x, y\}$ .

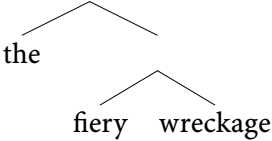
This would give to a string like (82) a representation like that in (83) on the next page.

<sup>11</sup> Our next reading: Chomsky (1995a).

## 8. Rethinking Things

---

(82) the fiery wreckage

(83)  =  $\{the \{fiery, wreckage\}\}$

One thing left out of (83) that seems necessary is that  $\{fiery, wreckage\}$  is a projection of *wreckage*, and that  $\{the \{fiery, wreckage\}\}$  is a projection of *the*. Chomsky suggests that we think of “projection” as a labeling of the sets formed by Merge. To represent that  $\{fiery, wreckage\}$  is a projection of *wreckage*, then, can be done by letting  $\{fiery, wreckage\}$  have *wreckage* as its label. Chomsky proposes the unfortunate means of representing this in (84).

(84)  $\{wreckage, \{fiery, wreckage\}\}$

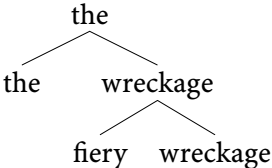
This doesn’t allow us to distinguish the result of MERGE from the result of labeling, however, so I suggest we adopt instead the notational convention in (85).

(85)  $\{wreckage \text{ } fiery, wreckage \}$

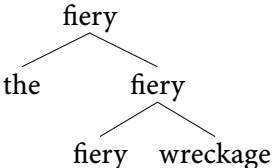
The notion of *projection*, then, can be expressed in terms of a labeling function. Modifying somewhat what Chomsky proposes, we can define this with (86).

(86)  $PROJECT(\{x,y\}) =_{\text{def.}} \{\alpha x,y\}, \alpha \text{ the LABEL}(\{x,y\}),$   
 where  $\begin{cases} \alpha = x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha = LABEL(x) \text{ or } LABEL(y) & \text{for } x, y \text{ phrases} \end{cases}$

Now, the representation that (82) receives is (87).

(87)  =  $\{the \text{ } the, \{wreckage \text{ } fiery, wreckage \} \}$

Of course, as far as PROJECT is concerned, (82) could also get the representation in (88).

(88)  =  $\{fiery \text{ } the, \{fiery \text{ } fiery, wreckage \} \}$

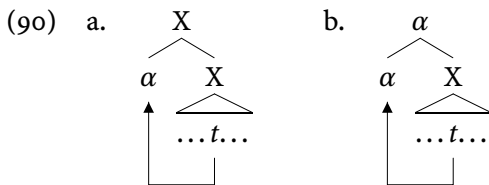
This representation, and others like it, will have to be blocked in the same way that such representations were blocked under the old system. That is, the category-specific information that is present in modification and c-selection blocks such cases. In (88), for instance, because *fiery* is an adjective, 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.

To let MERGE's effect be felt throughout the derivation, we must see it as part of the process that movement invokes. In this way the constraints on phrase structure rules invoked on D-structures will be spread through the other parses in a derivation. Let's understand movement, then, to be defined as in (89).

- (89) Move  $\alpha$  =<sub>def.</sub>
- a. MERGE  $\alpha$  to  $\beta$ , where  $\beta$  c-commands  $\alpha$ .
  - b. Insert a trace of  $\alpha$  in its former position, and interpret it as a variable bound to  $\alpha$ .

This defines all movement operations, including Argument Movement, as a kind of adjunction.

When MERGE is invoked in the movement operation, something different is needed of PROJECT in determining the label of the resulting phrase. In this scenario, we don't want the phrase that has moved to project, as in (90b), but instead the other phrase should, as in (90a).



What we look for is something that derives (91).

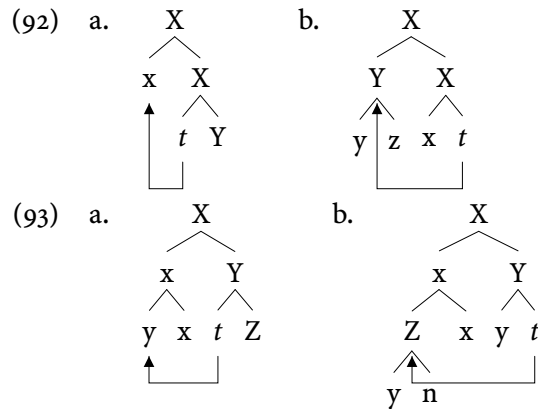
- (91) A moved term cannot PROJECT.

Let's assume that something — perhaps something yet to be built into the movement operation — derives (91).

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 (92) and (93).

## 8. Rethinking Things

---



In (92) are scenarios in which something has moved and Merged with a term that started out as a phrase; in (93) 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 (94).

- (94) An item cannot move from one position to another if it has a different projection status in those two positions.

This has the right consequences in those cases where movement adjoins something to a phrase: it permits (92b) and blocks (92a). But when movement adjoins something to a head, it badly misfires, blocking the expected outcome in (93a) but allowing the unexpected one in (93b).

In fact, the scenarios in (93) 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 (93) 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 (95).

- (95) THE WORD CRITERION  
If  $\text{MERGE}\{x, y\}$ ,  $x$  and  $y$  terminals, then Morphology must recognize  $\{x, y\}$  as a word.

To the extent that Morphology does not recognize phrases as words, this will block (93b). But in (93a), where  $x$  and  $y$  are terminals, Morphology is more likely to recognize the result as a word. Chomsky suggests that something along these lines — he doesn't propose (95) exactly — “overrides” (94) in scenarios like (93).

An interesting feature of these proposals, and one that Chomsky suggests we exploit, is that they allow derivations to mix structure building and movement. A derivation could start with a selection of terminal items — Chomsky calls this a NUMERATION or an ARRAY<sup>12</sup> — and assemble these terminals until a completed sentence, with all items moved, is derived. We can achieve this by changing our definition of Derivation to (96).

(96) Derivation

- a. Let  $\mathcal{R}$  be a transformational rule: that is, a function  $\mathcal{N}(N_i) = N_j$ , where  $N_i$  and  $N_j$  are numerations.
- b. A *Derivation* =<sub>def</sub> an ordered  $n$ -tuple of numerations,  $\mathcal{P} = (N_1, N_2, \dots, N_n)$ , such that:
  - i. Each  $N_i = \mathcal{R}(N_{i-1})$ , and
  - ii.  $N_1$  is a set of terminals.
  - iii. Some  $N_i, N_i$  a singleton, is phonologically interpreted.

(97) Instances of  $\mathcal{R}$ :

- i.  $\text{MERGE}(\alpha)(\beta) =_{\text{def}}$  replace two elements of  $N_i$ ,  $\alpha$  and  $\beta$ , with  $\{\gamma \alpha, \beta\}$ .
- ii.  $\text{MOVE}(\alpha)(\beta) =_{\text{def}}$  replace one element of  $N_i$ ,  $\alpha$ , with  $\text{MERGE}(\alpha')(\beta)$ , where  $\beta$  is contained in  $\alpha$  and  $\alpha'$  is just like  $\alpha$  except that in place of  $\beta$  is the trace of  $\beta$ .

On this definition, derivations are series of numerations, rather than series of phrase markers.<sup>13</sup> The initial numeration contains all the terminals to be used in constructing the sentence, and each subsequent numeration contains these formatives collected into increasingly more inclusive phrases.

What we've lost by jettisoning  $D$ -structure are all the constraints imposed on it. One of these is the restrictions imposed by the  $\bar{X}$  Skeleton; but this is recaptured by the LCA and MERGE. The others are the Projection Principle and the Modification rule, which, together, guaranteed that modifiers and arguments were placed correctly into a phrase marker. The Modification Rule is gone forever; let's hope that

<sup>12</sup> He has slightly different uses for these two terms; but what distinguishes them won't matter for us at this moment.

<sup>13</sup> This is not how Chomsky used the term “numeration,” and much of the literature follows his usage. He reserved that term for the initial collection of terminals.

## 8. Rethinking Things

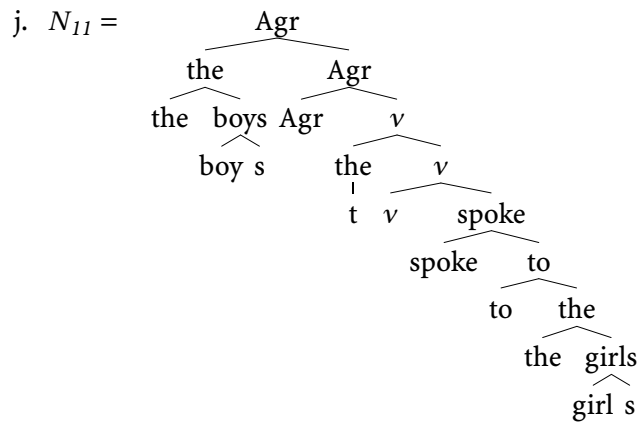
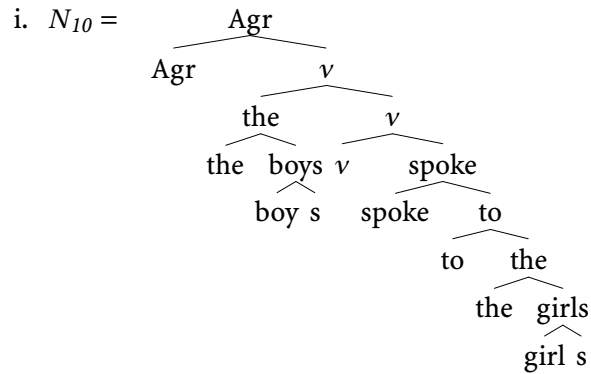
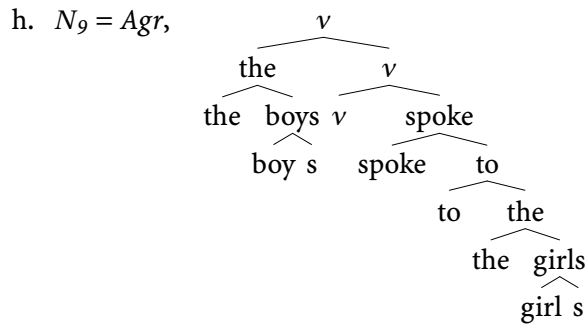
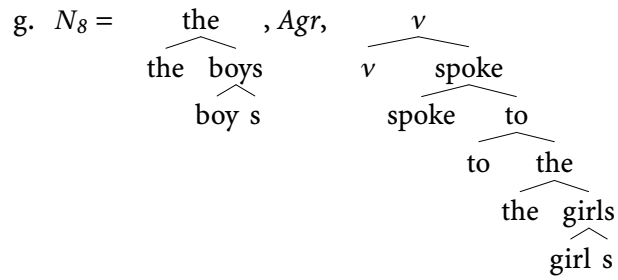
---

Cinque's project will take up that burden. The Projection Principle I will reintroduce in a moment. But first let's see how this definition of derivation will work.

Consider how the surface form for *The boys spoke to the girls* might look like (98).<sup>14</sup>

- (98)  $N_i = Agr, the, s, boy, v, spoke, to, the, s, girl$
- a.  $N_2 =$   $\begin{array}{c} \text{boys, Agr, the, v, spoke, to, the, s, girl} \\ \text{boys} \end{array}$
- b.  $N_3 =$   $\begin{array}{c} \text{the, Agr, v, spoke, to, the, s, girl} \\ \text{the boys} \\ \text{boys} \end{array}$
- c.  $N_4 =$   $\begin{array}{c} \text{the, Agr, v, spoke, to, the, girls} \\ \text{the boys} \quad \text{girls} \\ \text{boys} \quad \text{girl s} \end{array}$
- d.  $N_5 =$   $\begin{array}{c} \text{the, Agr, v, spoke, to, the} \\ \text{the boys} \quad \text{the girls} \\ \text{boys} \quad \text{girl s} \end{array}$
- e.  $N_6 =$   $\begin{array}{c} \text{the, Agr, v, spoke, to} \\ \text{the boys} \quad \text{to the} \\ \text{boys} \quad \text{the girls} \\ \quad \quad \text{girl s} \end{array}$
- f.  $N_7 =$   $\begin{array}{c} \text{the, Agr, v, spoke} \\ \text{the boys} \quad \text{spoke to} \\ \text{boys} \quad \text{to the} \\ \quad \quad \text{the girls} \\ \quad \quad \text{girl s} \end{array}$

<sup>14</sup> I've represented the results of MERGE in familiar "tree" notation; but understand these to be the structured sets that were defined above. The linear ordering that the trees represent should be understood as a consequence of the linearization algorithm.



## 8. Rethinking Things

---

Note that this derivation begins by building NPs up from the noun and the plural suffix. The definition of phrase we're working with has the funny consequence that every phrase "begins" by bringing two terminals together. When this is coupled with the WORD CRITERION, it forces derivations to start by bringing two morphemes together into a single word. Because I doubt this can be maintained, I suggest weakening the conditions on MERGING something with a terminal so that either the WORD CRITERION holds, or the Projection Principle does:

- (99) If  $\text{MERGE}(x)(y)$ ,  $x$  a terminal, then
- a.  $\{x, y\}$  must be recognized as a word by Morphology, or
  - b.  $x$  c-selects  $y$ .

This would allow for the derivation sketched in (98) to put number morphology together with the noun without making them a single word (under the assumption that number morphology c-selects nouns).

But (99) will also do some other work. It will capture the Likes-Attracts-Likes Constraint as well. We won't need Chomsky's condition on uniformity of "projection level" across derivations (i.e., (94)). And, it will prevent heads from being introduced in Specifier positions, which, recall, is something that the LCA failed to prevent unaided.

This doesn't completely capture the effects of the Projection Principle, however. It will require that a lexical item MERGE first with the term it c-selects (unless it can form a word with that term), but it doesn't require that all arguments be MERGED into a representation before modifiers start being introduced. To achieve this effect we must reinvokethe Projection Principle, but in a way that fits with the current model of derivations. Something like (100) will do.

- (100) For  $\text{MERGE}(x)(y)$ ,  $x$  and  $y$  phrases, if  $x$  has an unassigned  $\theta$ -role, then  $y$  must bear that  $\theta$ -role.

It's a live project to derive (100).

How do these proposals interact with Kayne's LCA? To reduce our old  $\bar{X}$  Theory to MERGE requires that the LCA, or something else, do the work of ensuring endocentricity. We'd like the LCA, then, to work smoothly with Chomsky's reduction of the phrase/head distinction. There are incompatibilities, however, and Chomsky sketches some ways to overcome them. As just noted, Chomsky's scheme requires that the first terminals that are put into a phrase always take the form in (101).

- (101) 
$$\begin{array}{c} x \\ \wedge \\ x \quad y \end{array}$$

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, and must



hold of every member of a derivation, then it will be dramatically non-total in this case, and this phrase marker should be ungrammatical. To allow for phrases of this sort, Chomsky suggests changing the criterion of Totality so that it holds only of phonetically interpreted terminals.

- (102) For all distinct, phonetically overt terminals,  $x$  and  $y$ , in a phrase marker, either  $[x,y]$  or  $[y,x]$ .

This will allow (101) just in case one of  $x$  or  $y$  is phonetically null. Another, similar, way of allowing (101) to survive the LCA by exploiting (102) is by letting  $x$  or  $y$  be a phonetically null terminal to begin with as, for instance, are the functional heads holding agreement, tense or number features. In this situation too, Chomsky's proposal in (102) would allow the LCA to linearize the overt member of  $x$  or  $y$ . A final way (102) could be exploited is to abandon Kayne's view that the LCA must hold of every phrase marker in a derivation, and let the LCA hold of the representation that arises 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. This will require, then, taking the LCA to hold of only certain phrase markers. Chomsky suggests the rather natural hypothesis that the LCA holds only of the pronounced phrase marker. This would let the contribution that the LCA makes to the shape of phrase markers only be felt on the "s-structure" parse. In principle, then, it's conceivable that the other phrase markers could avoid these consequences. Because those consequences play a role in deriving the Likes-Attracts-Likes Constraint — as we will see in the following section — I won't adopt this suggestion. As in Kayne's work, let's let the LCA hold of every phrase marker in a derivation.

Finally, we might imagine that the LCA is blocked from applying if  $x$  and  $y$  in (101) together form one word. In cases where MERGE brings together two subword parts, we might sensibly task the principles of Morphology with arranging those subparts. This could be achieved by restricting Totality even further, so that it applies only to words.

- (103) For all distinct, phonetically overt words,  $x$  and  $y$ , in a phrase marker, either  $[x,y]$  or  $[y,x]$ .

Even with these weakenings of Totality, combining the LCA with Chomsky's proposals in *Bare Phrase Structure* puts severe limitations on which terminals MERGE can combine. I do not know if these limitations are problematic.

Consider next the scenario in (104).

- (104)
- $$\begin{array}{c}
 x \\
 \wedge \\
 y \quad x \\
 \wedge \quad \wedge \\
 y \quad z \quad x \quad n
 \end{array}$$

## 8. Rethinking Things

---

This is a situation that presents a problem for the LCA, even without putting it together with Chomsky's proposals. Kayne labored mightily to overcome this problem, and, as we saw in the last section, he strove 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, 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*  $\{y, y, z\}$ . In this manner, mutual *c-command* of  $\{y, z\}$  and  $\{x, n\}$  is avoided and the terminals in these phrases will be linearized.

Chomsky proposes a similar strategy for (104), though he wishes to preserve the distinction between "adjuncts" and "specifiers" that Kayne dispenses with. Let's consider first, then, how Chomsky proposes to distinguish adjuncts and specifiers.

Chomsky takes the representation in (104) to put  $\{y, y, z\}$  in a Specifier position. For the LCA to apply with the outcome Kayne desires for these cases, Chomsky suggests restricting *c-command* with (105).<sup>15</sup>

(105) Only minimal and maximal projections can *c-command*.

This is very similar to the proposal I made in the previous section. However, whereas Chomsky, like Kayne, suggests blinding *c-command* to certain phrases so that the desired consequences arise, I suggested that the linearization algorithm should be blinded.

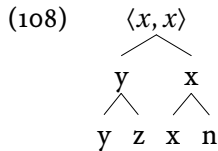
Adjunction structures, he suggests, have a different representation. He suggests that an adjunction structure differs from the way the Specifier is represented in (104) by virtue of the label on the phrase that Merge forms. In (104), the label on the phrase that contains the Specifier is a simple projection of the phrase that the Specifier is Merged with. In the case of adjunction structures, however, Chomsky proposes to give the phrase that contains the adjunct a special label. If the label of the node dominating a Specifier would be, say ' $x$ ', then the label of that node dominating an adjunct Chomsky suggests is  $\langle x, x \rangle$ . There are, in other words, two possible outcomes for PROJECT, the procedure that gives a phrase its label. This new PROJECT is defined in (106) on the facing page. To get the LCA to apply correctly to these structures, Chomsky suggests that phrases with the  $\langle x, y \rangle$  label dominate only the phrase that they are projected from. To see how these assumptions operate, consider the adjunction environment in (108), and adopt the definition of *c-command* in (107).

<sup>15</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.

(106)  $\text{PROJECT}(\{x,y\}) =_{\text{def.}} \{\alpha x,y\}$ ,  $\alpha$  the  $\text{LABEL}(\{x,y\})$ ,

$$\text{where } \left\{ \begin{array}{ll} \alpha = x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha = \text{LABEL}(x) \text{ or } \text{LABEL}(y) & \text{for } x, y \text{ phrases, or} \\ \alpha = \langle x, x \rangle \text{ or } \langle y, y \rangle & \text{for } x,y \text{ terminals, or} \\ \alpha = \begin{array}{l} \langle \text{LABEL}(x), \text{LABEL}(x) \rangle \\ \text{or} \\ \langle \text{LABEL}(y), \text{LABEL}(y) \rangle \end{array} & \text{for } x,y \text{ phrases} \end{array} \right.$$

(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 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 (101) 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.

In what follows I shall adopt both the LCA and the *Bare Phrase Structure* hypotheses. Because of the relative plasticity of the outcomes these hypotheses have for phrase structure, however, I 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  $\bar{X}$  theory that these proposals allow. Let us abandon the  $\bar{X}$  Skeleton, then, and adopt in its place MERGE, with the consequences that the LCA imposes on the resulting spoken parses. Concretely, let's adopt the following conventions:

## 8. Rethinking Things

---

- (109)
- a. The  $XP/X^0$  distinction is just the difference between a word and a group of words.
  - b. No distinction between Specifiers and Adjuncts.
  - c. Multiple adjuncts/Specifiers are permitted.
  - d. Terminals will be represented under  $X^0$ s.
  - e. Phrases will be given the traditional labels, e.g., “DP” “VP,” etc.

Representations with the properties of (109a-c) are consistent with the system we entertained as a variation of Kayne’s at the end of section 9.2. And I shall view the notational conventions in (109d) and (109e) 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.

### 8.3 Embedding the theory in a framework without $\bar{X}$ Skeleta

Let’s consider now how the theory we’ve developed will work with the ideas concerning phrase structure that we’ve just encountered. Our definition of derivations now takes us from a collection of terminals to a completely structured arrangement of those terminals. Derivations are series of Numerations, the name given to these increasingly structured collections of terminals.

(110) Derivation

- a. Let  $\mathcal{R}$  be a transformational rule: that is, a function  $\mathcal{R}(N_i) = N_j$ , where  $N_i$  and  $N_j$  are numerations.
- b. A *Derivation* =<sub>def</sub> an ordered n-tuple of numerations,  $D = (N_1, N_2, \dots, N_n)$ , such that:
  - i. Each  $N_i = \mathcal{R}(N_{i-1})$ , and
  - ii.  $N_1$  is a set of terminals.
  - iii. Some  $N_i, N_i$  a singleton, is phonologically interpreted.

(111) PROJECT( $\{x,y\}$ ) =<sub>def</sub>  $\{\alpha x,y\}$ ,  $\alpha$  the LABEL( $\{x,y\}$ ),

where  $\begin{cases} \alpha = x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha = \text{LABEL}(x) \text{ or LABEL}(y) & \text{for } x, y \text{ phrases} \end{cases}$

(112) Instances of  $\mathcal{R}$ :

- a.  $\text{MERGE}(\alpha)(\beta) =_{\text{def.}}$  replace  $\alpha, \beta \in N_i$  with  $\{\alpha, \beta\}$ .
- b.  $\text{MOVE}(\alpha)(\beta) =_{\text{def.}}$ 
  - i. replace  $\alpha \in N_i$  with  $\text{MERGE}(\alpha')(\beta)$ , where  $\beta$  is contained in  $\alpha$  and  $\alpha'$  is just like  $\alpha$  except that in place of  $\beta$  is the trace of  $\beta$ .
  - ii. replace  $\alpha \in N_i$  with  $\alpha'$ , where  $\alpha'$  is just like  $\alpha$  except that in place of some daughter of  $\alpha, \gamma$ , there is  $\text{MERGE}(\gamma)(\beta)$ ,  $\beta$  contained in  $\alpha$ , and in place of  $\beta$  there is the trace of  $\beta$ .

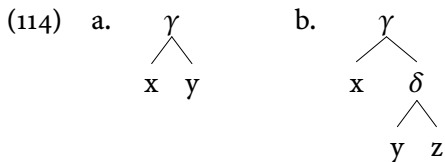
The two instances of  $\text{MOVE}$  correspond to Head Movement ( $=$ (112b.ii)) and Argument Movement ( $=$ (112b.i)).

The Likes-Attracts-Likes Constraint and the Word Criterion are now derived from the following constraint on  $\text{MERGE}$ .

- (113) If  $\text{MERGE}(x)(y)$ ,  $x$  a terminal, then
- a.  $\{x, y\}$  must be recognized as a word by Morphology, or
  - b.  $x$  c-selects  $y$ .

Because  $\text{MERGE}$  plays a role both in building up a phrase marker and in moving things around, the effects of (113) will be felt in both applications. Let's take a moment to review how (113) derives these former principles. Consider the two possible ways in which a terminal might be merged into a phrase marker.

First imagine how a terminal might be merged into a phrase marker as an independent member of the Numeration. (This is a situation in which the terminal is not moving.) This will form one of the two structures in (114), depending on whether it is merged with a phrase or another terminal.

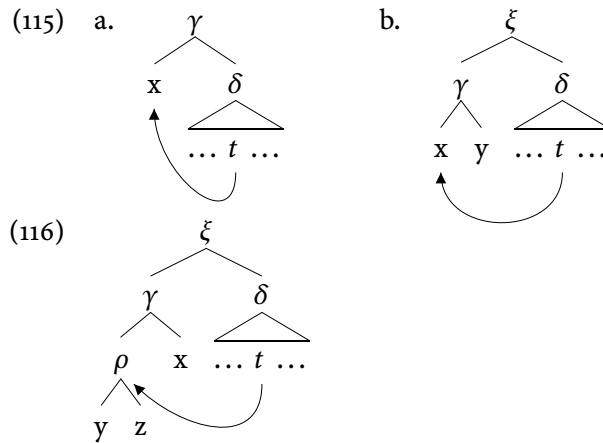


In (114a),  $\gamma$  will either be a word, or  $y$  will be an argument of  $x$ , and  $\gamma$  will be a phrase. We've not yet encountered scenarios where either of these things happen; but we've also not seen any reason to block them. Let's leave this open, then. In (114b), either  $x$  and  $\delta$  will have to form a word, or  $\delta$  will be  $x$ 's complement. We've not seen cases of the first sort, but we've seen plenty of examples like the latter. Essentially, then, in building up a phrase-marker that brings a head and something else together, (113) permits what we've seen we need (i.e., head and complement becoming sisters) and also ways of forming complex words from heads.

## 8. Rethinking Things

---

Consider next how a terminal might be merged into a phrase marker as part of a movement operation. Here we have two cases to consider. In one (=115), the terminal is moved; and in the other (=116) a phrase is moved.



The case in (115b) is the one we've seen evidence for: one head adjoining to another. This is permitted by (113) just in case  $\gamma$  is recognized as a word. In this way, then, the Word Criterion's effects are captured. The case in (115a) arises when a head adjoins to a phrase, and this is what the Likes-Attracts-Likes constraint was designed to prevent. It will be allowed by (113), but only if  $x$  and  $\delta$  can form a word or if  $\delta$  is c-selected by  $x$ . We don't know if the scenario in which  $x$  and  $\delta$  form a word should be disallowed. This would be a case in which movement allows a part of a word to be rearranged within that word. The Likes-Attracts-Likes Constraint was not designed for this case, and so it would have allowed it as well. Movement within the word level has sometimes been entertained.<sup>16</sup> We might even imagine that particle verbs instantiate something of this sort. Let's leave this possibility open.

If  $\delta$  does not form a word with  $x$ , however, it will have to be c-selected by  $x$ , and this is plausibly blocked. Consider how such an arrangement would arise. Before  $x$  moves, it will be a head within  $\delta$ . If  $x$  combines with a phrase within  $\delta$  it will be required to either form a word with that phrase or assign a  $\theta$ -role to it. Let's call this subphrase — the one that is formed by merging  $x$  within something inside  $\delta$ :  $\alpha$ . There must be a  $\theta$ -role left-over for  $x$  to assign once  $\alpha$  is constructed, or otherwise  $x$  will not be able to take  $\delta$  as an argument.<sup>17</sup> But if there is a  $\theta$ -role left-over upon completion of building  $\alpha$ , then the Projection Principle, whose new form is (117), will require the phrase that  $\alpha$  merges with to get that  $\theta$ -role.

<sup>16</sup> See Pesetsky (1985) for an example.

<sup>17</sup> This follows on the assumption that a  $\theta$ -role may not be assigned twice, and this is what the Theta Criterion requires.

- (117) For  $\text{MERGE}(x)(y)$ ,  $x$  and  $y$  phrases, if  $x$  has an unassigned  $\theta$ -role, then  $y$  must bear that  $\theta$ -role.

As a consequence, (113) will allow (115a), but only if  $\delta$  is made up of just  $x$  and at most one complement of  $x$ . That is so limited a scope for movement of a head to a phrase, that it may do no harm. In other words, it derives enough of the consequences of the Likes-Attracts-Likes Constraint that it may be empirically equivalent to it.

Finally, consider the situation in (116). In this scenario, a phrase has moved and adjoined to a head. This too is an outcome blocked by the Likes-Attracts-Likes Constraint. It will be allowed by (113) only if the phrase forms a word with the head, or if the phrase is  $c$ -selected by the head. There is no reason to block the case where the phrase forms a word with the head since this amounts to a situation in which one word adjoins to another, and the Likes-Attracts-Likes Constraint would have allowed that to begin with. If the phrase that moves doesn't form a word with  $x$ , however, then it must be an argument. Because (113) requires this phrase to be an argument of  $x$ , the position the phrase has moved from must be a non- $\theta$ -marked position.<sup>18</sup> But these derivations plausibly violate EARLINESS. They will involve derivations where arguments are introduced in places that they are not interpreted in and then moved into places where they are. Such derivations will be longer than ones in which the arguments are merged directly into the positions where they get their  $\theta$ -role, and so EARLINESS will rule them out.

Together, then, these constraints construct binary branching phrase markers, in which heads merge first with their arguments and the phrases they build then combine with modifiers, possibly, and then form arguments for other heads. Things can move around in constructing these phrase markers, and since there are only two things – heads and phrases – only heads and phrases will move. The places these things can move to are constrained in the way just described, which closely mimics what the Likes-Attracts-Likes Constraint and Word Criterion achieved.

In addition to (113) and (117), there are constraints on derivations that govern how far MOVE may relocate an item. For Argument Movement that locality condition is the modified Specified Subject Constraint; for Head Movement it is the Head Movement Constraint. Our working hypothesis is that these two can be collapsed into a single constraint: Relativized Minimality.

- (118) ALMOST RELATIVIZED MINIMALITY  
 $X^n$  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.

<sup>18</sup> This follows from the Theta Criterion.

This constraint — or whatever turns out to be the right family of locality constraints — will determine how many steps in the derivation are necessary to achieve a pronounceable representation. If it is not possible to move an item from its initial position in a phrase marker directly into the position it which it can be pronounced because of these locality conditions, then derivation will include Numerations in which that time has moved to intermediary positions (or there will be no pronounceable outcome). Derivations of this sort are said to involve “successive cyclic movement,” of the item in question.

From any given initial Numeration, it may be possible to construct a pronounceable representation in a variety of ways. If there are more than one such derivations, EARLINESS will, as before, choose that derivation that brings about this representation in the fewest number of steps. EARLINESS will have to take a slightly different form, now, because derivations are now series of Numerations, rather than series of phrase markers. EARLINESS can be formulated as in (119).

(119) EARLINESS

Let  $\mathcal{D} = \{D_1, D_2, \dots, D_n\}$  be the set of well-formed derivations for some sentence,  $S$ , and  $\mathcal{N}$  be the set of  $n$ 's, such that for every  $D_i \in \mathcal{D}$ ,  $n_i$  is that sub-series of numerations in  $D_i$  that starts with the initial numeration of  $D_i$  and ends with the first pronounceable numeration in  $D_i$ . The pronounced numeration of  $S$  is the one in the shortest  $n \in \mathcal{N}$ .

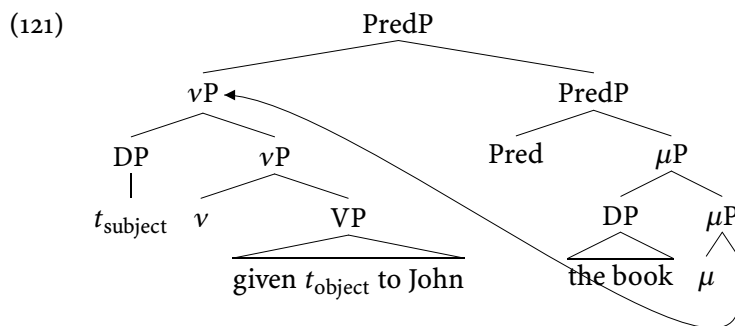
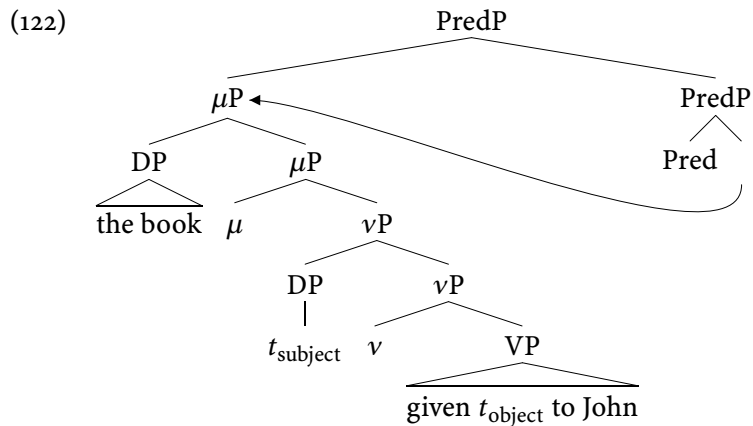
EARLINESS, it will be recalled, was introduced to prevent main verbs from moving overtly. Because  $\text{Agr}^0$  has weak features, movement of a verb to check those features is not required to make a pronounceable representation and EARLINESS therefore prevents this movement from forming the pronounceable parse. We've encountered other uses for EARLINESS along the way. It may play a role in preventing phrases from moving to head positions, as we've just seen. And it could help constrain the places where quantifiers float, as we saw in the previous chapter.

But there are also problems for EARLINESS. One of these that was discussed in the previous chapter is that it seems to restrict too narrowly the places where quantifiers float. It blocks, for example, the derivation that would be necessary to derive (120).

(120) The kids have all been dressed.

It may also seem inconsistent with Heavy NP Shift, which, at least in the form it takes in John Ross's work, is an optional rule. It is worth noting, however, that on the view of Heavy NP Shift sketched in section 8.2.1 there is no conflict with EARLINESS. On that account, Heavy NP Shift word orders arise by moving a predicate leftwards that does not include the shifted term, as indicated in (121).





The non-Heavy NP Shifted word orders are arrived at on this view by moving instead the entire  $\mu\text{P}$ , as in (122) below. In both derivations each step is obligatory and so in neither case is a shorter derivation possible. The difference in the outcomes hinges on the variability of the target of Predicate Shift, and not on varying the derivations. If this account of Heavy NP Shift can be generalized to all such cases, then it will be possible to square these word order alternations with EARLINESS. This would leave only the cases of Quantifier Float in (120) to be reconciled.

*Earliness* has a much more general utility than the few cases its been put to use for so far. It prevents surface representations from diverging too dramatically from their underlying form.<sup>19</sup> It allows terms to move from the positions they are originally MERGED into only when not doing so would fail to meet the requirements of pronounceability. In the core set of cases we have been examining, CPs and DPs move to check off Case features, and verbs, nouns and the functional heads they

<sup>19</sup> It has the same function that faithfulness constraints do in Optimality Theory. Like *Earliness*, faithfulness constraints pressure surface forms to conform to underlying forms as much as possible.

## 8. Rethinking Things

---

combine with do the same. All other instances of movement will be blocked by EARLINESS. It will prevent cases such as (123), for instance, where an NP has moved out of a DP; and it will prevent movement of AdvPs, AgrPs, and so on.

(123) \* Cat seems <sub>[DP the t]</sub> unhappy.

EARLINESS is what prevents MOVE from scrambling sentences to pieces. It forces the parse that is pronounced to be as close as possible to the one that is constructed by MERGE, allowing divergence only to the extent that meeting the conditions on pronounceability requires.

A final constraint on derivations is the linearization algorithm, which places restrictions on what sorts of phrase markers each application of  $\mathcal{R}$  may form. Our linearization algorithm is a modification of Kayne's Linear Correspondence Algorithm. It relies on the following well-formedness constraints on a linearization.

- (124) Let  $x, y$  and  $z$  be terminal words, and “[ $\alpha, \beta$ ]” mean that  $\alpha$  linearly precedes  $\beta$ .
- If [ $x, y$ ] and [ $y, z$ ], then [ $x, z$ ]. (it's Transitive).
  - For all distinct, phonetically overt,  $x$  and  $y$  in a phrase marker, then either [ $x, y$ ] or [ $y, x$ ]. (it's total).
  - not ([ $x, y$ ] and [ $y, x$ ]) (it's Antisymmetric).

A linearization, then, is defined in (125)

- (125)
- $\alpha$  c-commands  $\beta$  iff every phrase dominating  $\alpha$  dominates  $\beta$   $\alpha$  excludes  $\beta$  and  $\beta$  does not dominate  $\alpha$ .
  - $\alpha$  asymmetrically c-commands  $\beta$  iff  $\alpha$  c-commands  $\beta$ , and  $\beta$  does not c-command  $\alpha$ .
  - $d(X) =_{\text{def}}$  the set of all terminals dominated by  $X$ .
  - $d(\langle X, Y \rangle) =_{\text{def}}$  the set of all ordered pairs  $\langle x, y \rangle$  such that  $x$  is dominated by  $X$  and  $y$  is dominated by  $Y$ .
  - $d(\langle X_1, Y_1 \rangle, \langle X_2, Y_2 \rangle, \dots, \langle X_n, Y_n \rangle) =_{\text{def}}$   $d(\langle X_1, Y_1 \rangle) \cup d(\langle X_2, Y_2 \rangle) \dots \cup d(\langle X_n, Y_n \rangle)$ .
  - Let  $A$  be the set of all  $\langle X, Y \rangle$  in some phrase marker such that  $X$  asymmetrically c-commands  $Y$ , where  $X, Y$  are categories.
  - $d(A)$  is a linearization.

This linearization allows multiple adjunct/specifiers, but places them always on the same side. We specify how the linearization is interpreted on a language, or perhaps phrase, specific basis with (126).

(126) For  $\alpha$  a head:  $\langle \alpha, \beta \rangle = [\alpha, \beta]$  or  $[\beta, \alpha]$ .

In all other cases, our hypothesis is that the linearization algorithm produces a “precedes” relation:

(127) For  $\alpha, \beta$  phrases:  $\langle \alpha, \beta \rangle = [\alpha, \beta]$ .

This corresponds roughly to our initial linearization typology. But it does so in a way that explains the correlation between being higher in a phrase marker and being linearized to the left.

This brings us to the final component of the grammar: the conditions on pronounceability. With the translation of the Case filter into feature terminology, and the casting of PRO’s surface position into these terms as well, our conditions on pronounceability all boil down to feature checking. The central criterion is (128).

(128) A term with an unchecked strong feature cannot be pronounced.

Features get checked when they are matched by features on another term and one of the two locality conditions in (129) hold.

(129) FEATURE CHECKING

- a. A feature  $[\phi]$  on a phrase,  $\alpha$ , is checked when  $\alpha$  is in the Specifier of  $\beta$  and  $\beta$  bears the matching feature  $[\phi]$ .
- b. A feature  $[\phi]$  on a head,  $\alpha$ , is checked when  $\alpha$  is adjoined to  $\beta$  and  $\beta$  bears the matching feature  $[\phi]$ .

These statements make use of notions from our former view of phrase markers, and can be collapsed to the simpler statement in (130), as we will see.

(130) FEATURE CHECKING

A feature  $[\phi]$  on  $\beta$  is checked when  $\text{MERGE}(\alpha)(\beta)$ , where  $\alpha$  bears a matching  $[\phi]$ .

To see how this works, and to fix some of the details, let’s consider how the derivation of the simple French sentence in (131) goes.

(131) Jean aime Marie.  
 John likes Mary  
 ‘John likes Mary.’

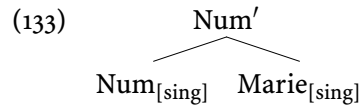
The initial Numeration might be (132).

(132)  $N_1 = \{D, \text{Num}_{[\text{sing}]}, \text{Jean}_{[\text{sing}]}, \text{Agr}_{[\text{3rd, sing}]}, T_{[\text{pres}]}, \mu_{[\text{acc}]}, v,$   
 $\text{aime}_{[\text{3rd, sing}]}, D, \text{Num}_{[\text{sing}]}, \text{Marie}_{[\text{sing}]}\}$   
 $\text{[pres]}$

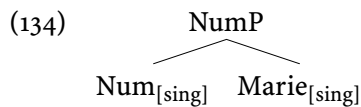
## 8. Rethinking Things

---

Imagine that MERGE first applies to *Marie* and Num, and replaces them with the structure in (133).

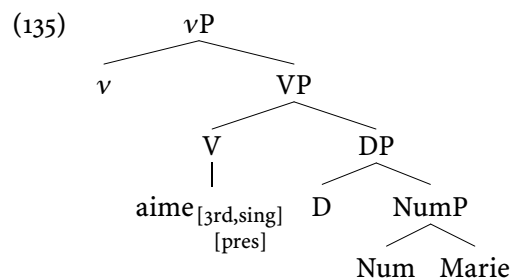


We have to decide whether Num' is interpreted as a word or a phrase. If it is a phrase, then either Num must *c*-select *Marie*, or *Marie* must *c*-select Num if it is to satisfy (113). I don't know how to decide between these options. Let's assume for concreteness that Num' is a phrase and that Num *c*-selects *Marie*. In keeping with the labeling conventions I invoked at the end of the last section, (133) would be represented as (134).



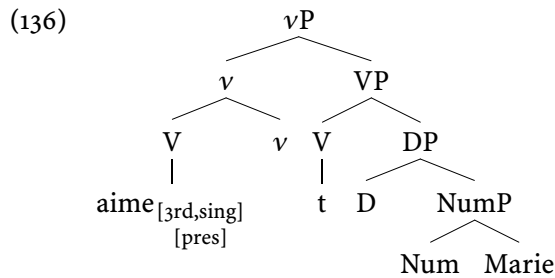
This structure meets the requirements of feature checking, and so the matched [sing] features will be checked. This is also a representation that can be linearized, since only one of the terminals here is phonetically overt.

Assuming that D, *aime* and *v* *c*-select NumP, DP and VP respectively, MERGE will be capable of applying iteratively to replace these terms with (135).



I hereby adopt the convention of removing features from the syntactic representation when they have been checked. One of the MOVE operations (viz. (112b.ii)) can now apply to adjoin *aime* to *v*. Because this rule is restricted so that it can only adjoin something to a daughter of a member of a Numeration, it will not be able to apply later in the derivation.<sup>20</sup> I've imposed this constraint arbitrarily — we've not encountered any evidence for it — merely to make the derivations slightly more deterministic and therefore more tractable to describe. From (135), then, will be formed (136).

<sup>20</sup> At this stage in the derivation, (135) is one of the members of the Numeration, but as soon as something MERGES with (135) it won't be.

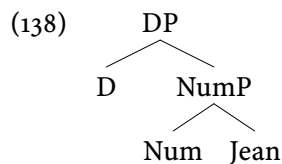


At this point in the derivation, the Numeration will contain (136) and the terms in (137).

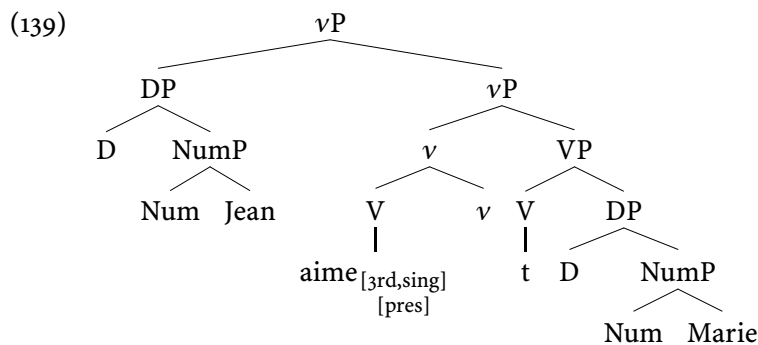
(137)  $N_4 = \{D, \text{Num}_{[\text{sing}]}, \text{Jean}_{[\text{sing}]}, \text{Agr}_{[\text{3rd, sing}], T}_{[\text{pres}]}, \mu_{[\text{acc}]}\}$

$\begin{matrix} [\text{nom}] \\ [\text{EPP}] \end{matrix}$

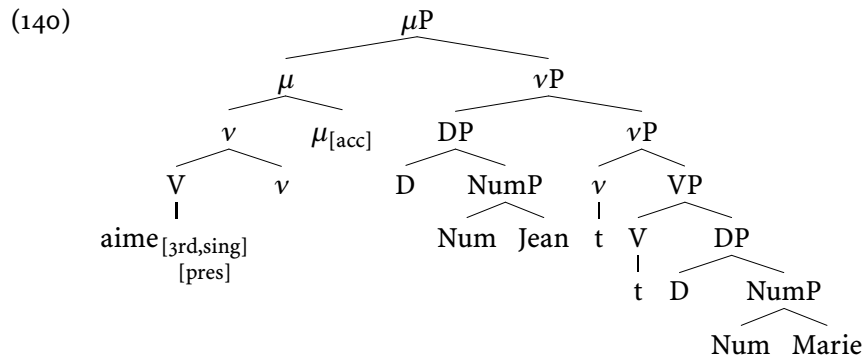
We cannot now MERGE  $\mu$  to (136) because  $v$  has a  $\theta$ -role to assign, and (117) requires that the next thing that MERGES with  $vP$  get that  $\theta$ -role. So at this stage MERGE must construct the DP that is that argument. It will therefore apply recursively to  $D$ ,  $\text{Num}$  and  $\text{Jean}$  to form the DP in (138).



Now MERGE can bring these two phrases together to form (139).



Now it is possible to MERGE  $\mu$ , and this will be followed by an instance of MOVE to form (140).



The next step in this derivation should move the object DP and merge it to  $\mu$ P. This movement is necessary to check off the [acc] feature on  $\mu$ . But as can be seen in (140), I have not given a matching [acc] feature on the object DP. The proposal is that all overt DPs have a overt Case feature, and that feature should be [acc] for the object DP in (140). We should now stop a moment and consider how to bring this feature into the derivation.

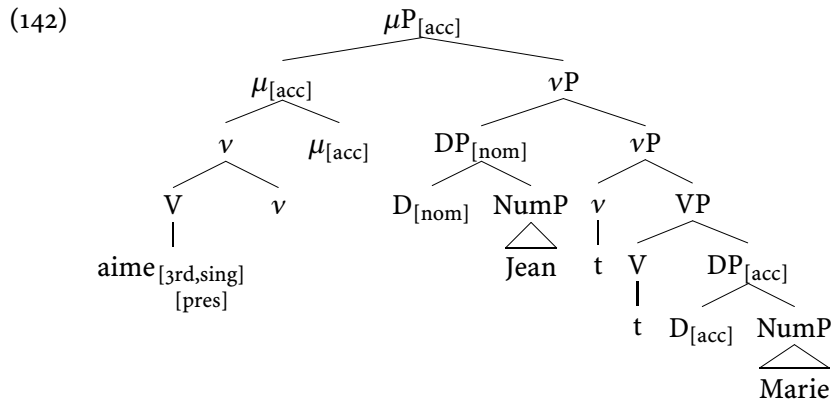
All the other features in this derivation are present on the terminals in the original Numeration. We can think of these features as parts of the lexical entries for these terminals. How should we think of features that are on phrases? Phrases, after all, are nothing more than labeled groupings of terminals in this framework. One idea is that there is a process that applies to certain phrases and adds features to them. We might think of this as another member of the set of  $\mathcal{R}$ 's that generate new members of a derivation. Alternatively, we might imagine that these features are present on some lexical item in the original Numeration and are transferred, somehow, to a phrase. Chomsky has suggested the second alternative, and has formulated a constraint that would ensure this. He calls this constraint INCLUSIVENESS, and it requires that all of the material that go into making a sentence — features, lexical item, what have you — are present in the initial Numeration. I'll follow that route here too.

In this case, let us imagine that it is the determiners that bring a Case feature. The initial Numeration for this sentence, then, should have had “D<sub>[acc]</sub>” and “D<sub>[nom]</sub>” instead of the two D's shown in (132). These features are strong; they are the sort that will make the things bearing them unpronounceable unless they become checked. Further, suppose it is the PROJECT relation that gives features borne by some terminal to a phrase that contains that terminal. The formulation of PROJECT should be changed from (111) to (141).

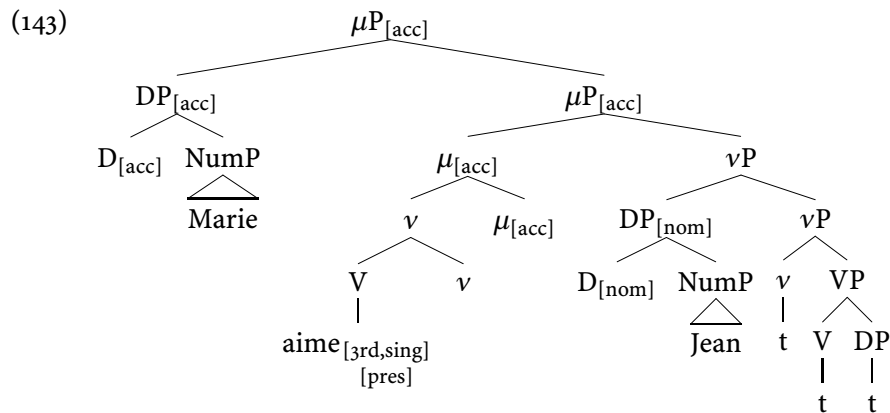
$$(141) \text{ PROJECT}(\{x,y\}) =_{\text{def.}} \{\alpha x,y\}, \alpha \text{ the LABEL}(\{x,y\}),$$

$$\text{where } \begin{cases} \alpha \text{ has the features of } x \text{ or } y & \text{for } x, y \text{ terminals, or} \\ \alpha \text{ has the features of LABEL}(x) \text{ OR LABEL}(y) & \text{for } x, y \text{ phrases} \end{cases}$$

These moves, then, would change the representation in (140) to one in which Case features are present on not just the lexical items, but all of their projections as well. (142) illustrates.



From (142) MOVE can construct (145).



This representation is the one at which the [acc] feature is to be checked. It is clear from the definition of feature checking (repeated here) that this configuration would check the [acc] on DP and the smaller  $\mu$ P.

(130) FEATURE CHECKING

A feature  $[\phi]$  on  $\beta$  is checked when  $\text{MERGE}(\alpha)(\beta)$ , where  $\alpha$  bears a matching  $[\phi]$ .

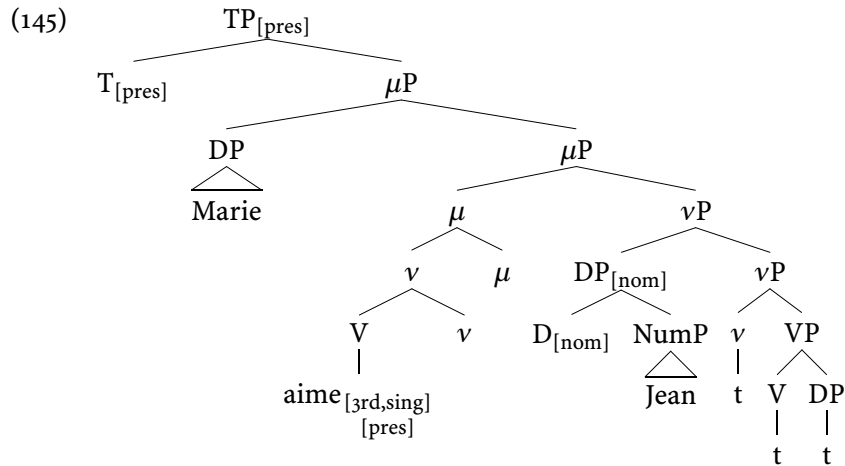
But we must also check the [acc] feature on D,  $\mu$  and the higher  $\mu$ P. Let us adopt the convention in (144).

(144) If a feature on  $\alpha$  PROJECTS from or PROJECTS to  $\beta$ , then checking  $\alpha$ 's feature checks  $\beta$ 's too.

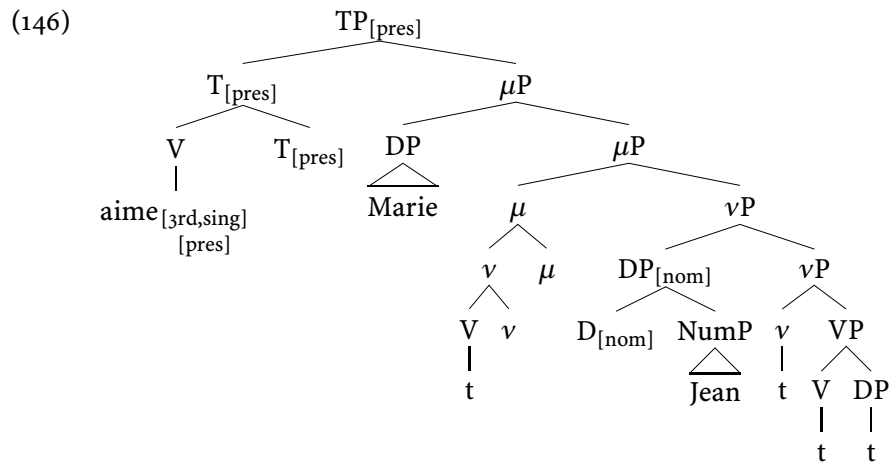
## 8. Rethinking Things

This will have the desired consequence of checking all the [acc] features in (142).

The next step in the derivation merges  $T_{\text{pres}}$ .



MOVE should apply to this representation to bring *aime* and T together. This step is necessary to check the [pres] features of T which could be strong (it's hard to tell in French finite clauses). In the derivations involving Head Movement that were considered earlier, Head Movement always affected just the highest head when a group of heads had merged.<sup>21</sup> If that occurred here, it would MERGE (the highest)  $\mu$  to T. That representation, however, would not allow (130) to check the [pres] feature. Instead, MOVE must merge *aime* to T forming (146).

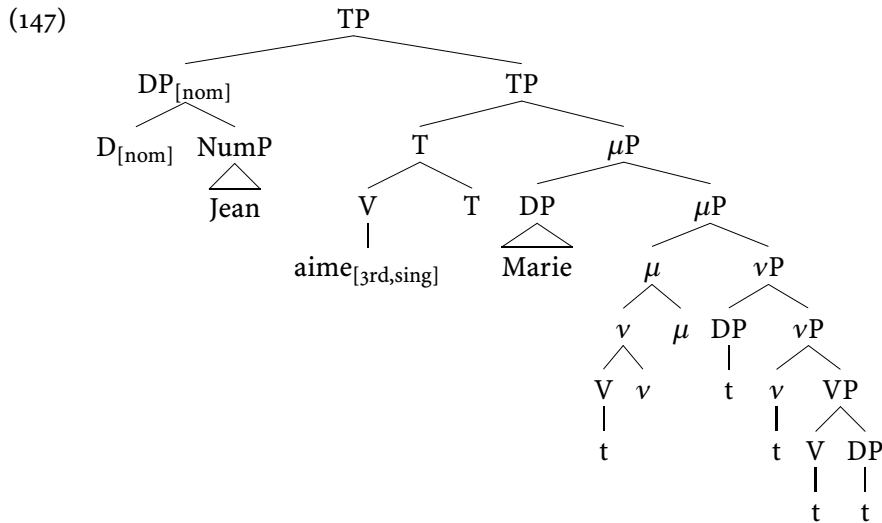


This is another demonstration that on this system, it is the placement of features that determines which things move. In (146) all the [pres] features become checked.

<sup>21</sup> With the notable possible exception of particle verbs.



MOVE must apply again, MERGING *Jean* to TP:



This step is made necessary by the locality condition on Argument Movement. That locality condition, recall, will not let *Jean* move past *Marie* unless it lands in a position that is in the same minimal domain as is *Marie*. When *aime* moves to T, it creates a minimal domain that includes phrases MERGED with TP and phrases merged with  $\mu$ P. After more material MERGES with (147), it will not be possible to for *Jean* to move into this minimal domain,<sup>22</sup> so it must do so at this point.

The next step MERGES the last remaining members of the Numeration: Agr and (147). This is followed by MOVE merging *aime* with Agr to check off the [3rd,sing] feature. This forms (148) on the following page. To this representation MOVE will again apply, bringing *Jean* into the position where its Case features can be checked, and this forms what will be the surface representation in (149) on the next page. This configuration checks the [nom] feature. It also is intended to check the [EPP] feature on the projections of Agr, however there is no matching feature in (149) that can play this role. The [EPP] feature is special: it is satisfied if any phrase comes into a checking relationship with it. Let's record this special promiscuity of [EPP] with (150).

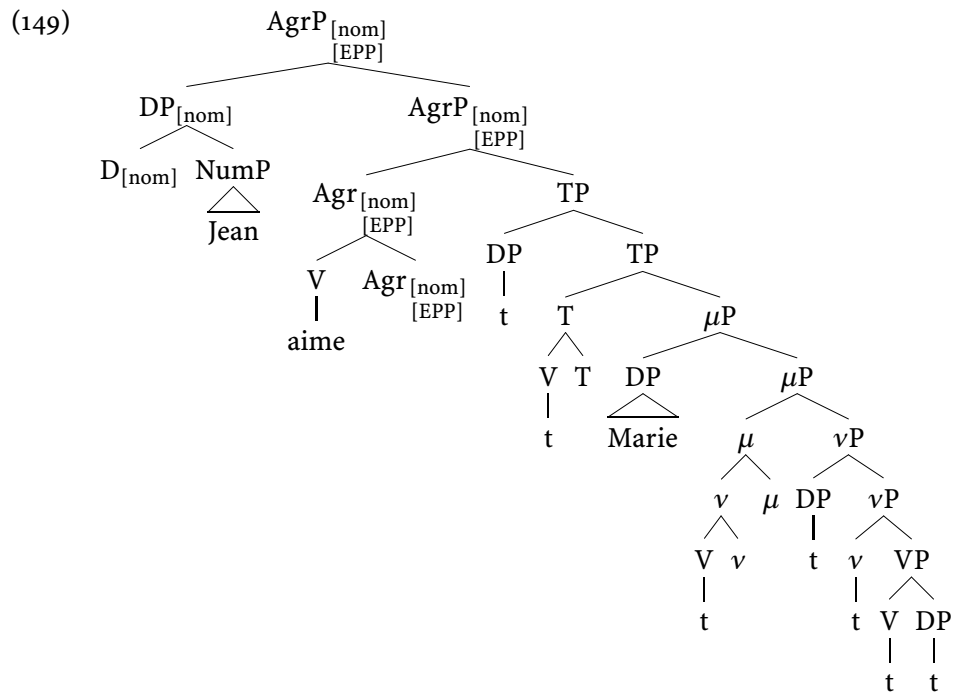
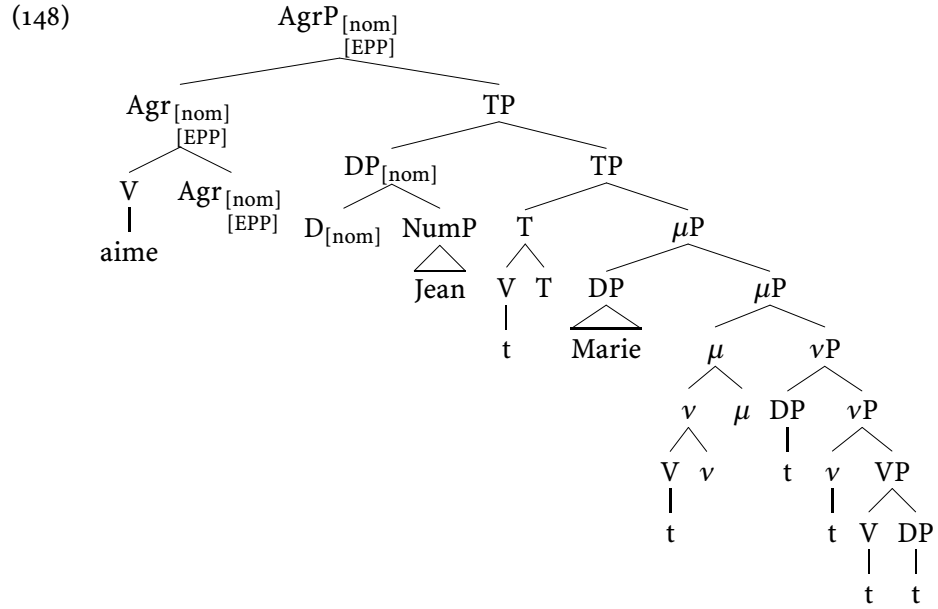
(150) The [EPP] feature matches any phrase.

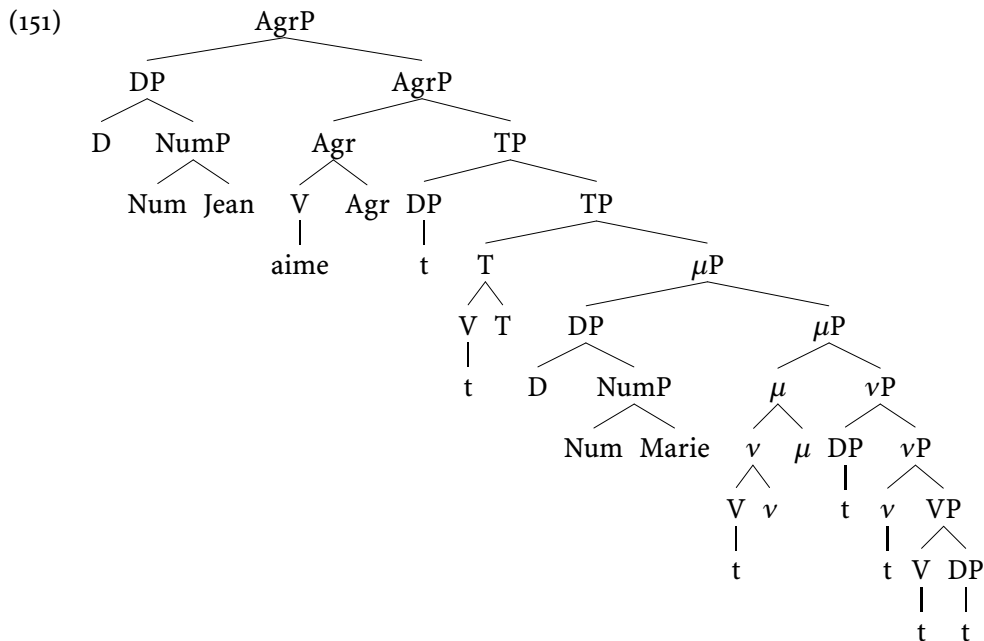
With this in mind, the representation in (149) becomes (151) on page 341.

<sup>22</sup> This follows from the way MOVE is designed when it applies to phrases (see (112b.i)). As with the statement of Head MOVEMENT, I have arbitrarily restricted Argument Movement so that it can only MERGE things to an element of the Numeration it applies to.

8. Rethinking Things

---





Not only are all the strong features checked in (151), all the features are checked. This, then, is the first step in the derivation that is pronounceable, and since this derivation is also the shortest to get to a pronounceable parse, this is the representation that will be pronounced. It is therefore subject to the linearization procedure, where the head parameter sets heads to precede their sisters. There is a successful linearization of this phrase marker, and it puts the words in the right order.

There are a couple of problematic aspects of this system to point out. One is that there is nothing that expresses the fact that the features on Agr and the verb must correspond with the number and person properties of the subject. That is there is nothing that expresses the fact that the verb must agree along these dimensions with the subject. All the [3rd,sing] feature presently does is express the fact that the morphology on the verb must correspond to the features on Agr. Something needs to be added to express agreement; let's adopt the descriptive (152) as a stopgap until this mechanism is discovered.

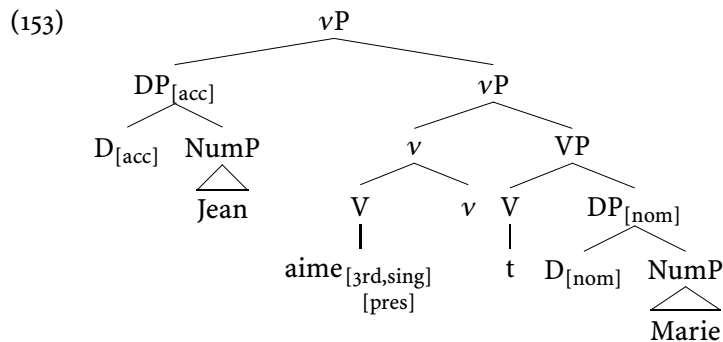
(152) AGREE

The person/number feature on Agr must agree with the person and number properties of the phrase that checks its Case feature.

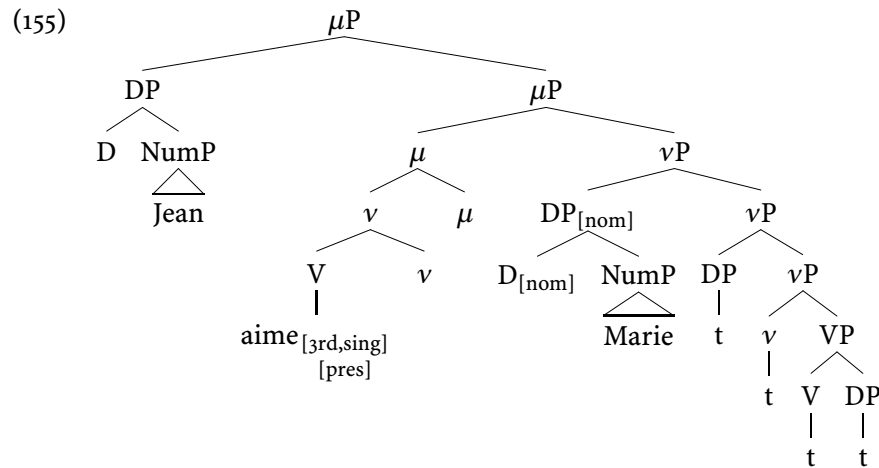
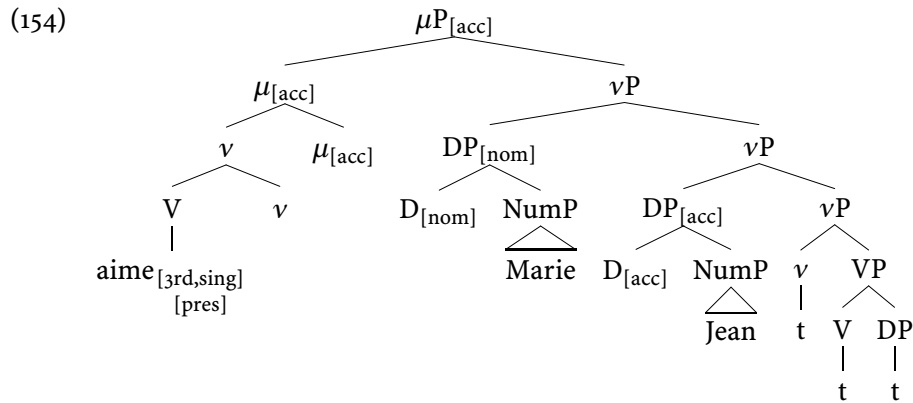
That AGREE should make reference to the Case feature records the fact that, in general, finite verbs agree with the Nominative Case marked DP. We should strive to

build this linkage into the feature system, no doubt.

There is also a problem with the locality condition on Argument Movement. Its effects do not survive the transition to a system of phrase markers, like that which we have adopted, in which the distinction between Specifier and Adjunct is removed and there can be an arbitrary number of Specifier/Adjuncts per phrase. In abandoning the  $\bar{X}$  Skeleton approach to constraining phrase markers, we lost the ability to limit the number of Specifiers per Phrase to just one. This made a crucial contribution to how the modified Specified Subject Constraint operated. Because Argument Movement was limited to moving phrases into Specifier positions, it was necessary to have an open Specifier position immediately above some intervening DP that was being crossed. On the present system, however, this effect is no longer derived. To see this, imagine that the subject and object DPs had been constructed with the ‘wrong’ Case features and consider the point in the previous derivation where the phrase in (153) has been constructed.



It will be possible, and necessary, for MOVE to MERGE the object DP to  $vP$  at this stage, after which  $\mu$  can be MERGED and  $v$  moved to  $\mu$ . The resulting representation is (154) on the next page. Moving the object DP to  $vP$  satisfies the locality condition on Argument Movement, as it puts it in a position that is in the same minimal domain as the subject DP it has passed. All of these steps satisfy the conditions on derivations. Because  $v$  has moved to  $\mu$ , there is now a minimal domain that includes the object DP and phrases that are merged to  $\mu P$ . As a consequence, it is now possible to move *Jean* to  $\mu P$ , as indicated in (155) on the facing page, where the [acc] features will be checked. From this representation it will be possible to MOVE the object DP, *Marie*, so that it MERGES with  $\mu P$ , and eventually makes its way to AgrP, where its [nom] feature will be checked. The availability of this derivation means that, in general, it is possible for object and subject DPs to move into the wrong Case marked positions. The locality condition on Argument Movement was intended to prevent this outcome; its effects have been undermined by our revisions to phrase markers.



To recapture these effects, we must either prevent more than one adjunct/Specifier per phrase, as on Kayne's proposals, or re-craft the locality condition. We will take up the latter direction in the following chapter.

But before turning to that task, let me spell out a popular variant of the system we've arrived at. This variant, found in Chomsky (1995b) and Kratzer (1996), collapses the distinction between  $\mu$  and  $v$ . At present, we express the dependency between Accusative Case assignment and the availability of a subject  $\theta$ -role that Burzio's Generalization describes with (156).

(156) Only  $\mu$  c-selects  $v$ .

On this variant, there is no  $\mu$  and this dependencies is expressed with (157).

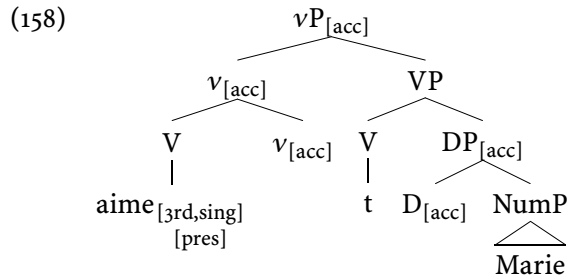
## 8. Rethinking Things

---

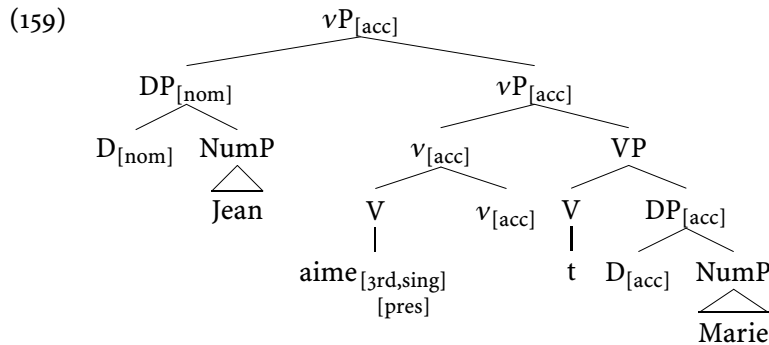
(157) Only  $v$  (and D) can have the [acc] feature.

On this variant, an accusative Case marked DP adjoins to  $vP$ , not  $\mu P$ .

To see how this works, consider the stage in the derivation of *Jean aime Marie* in which  $vP$  has been constructed.

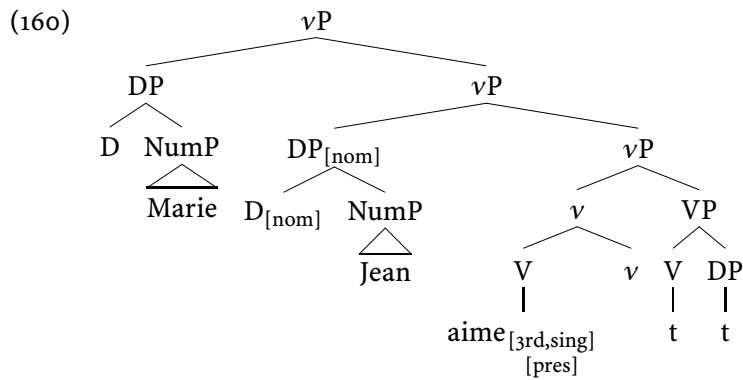


The next term that MERGES with  $vP$  must be the phrase that  $vP$  assigns a  $\theta$ -role to. This is what the remnant of the Projection Principle in (117) requires. For this reason, the object DP will not at this point adjoin to  $vP$ . From (158) is derived (159), then.



At this point, MOVE can *merge* the object DP to  $vP$ , thereby checking the [acc] features, as in indicated in (160) on the next page. From this point, the derivation will proceed as before, with *Jean* eventually MOVING to AgrP, where it will check off the [EPP] and [nom] features.

Notice that this variant ensures that the Accusative Case marked object surfaces higher than the underlying position of the subject argument, just as the original system does. (This outcome, recall, is necessary to capture the facts concerning *again*.) On the other system this is ensured by forcing  $\mu P$  to embed  $vP$ , which follows from (156). On this variant it emerges as a consequence of two constraints: (117), which forces the subject argument to be MERGED with  $vP$  before the object argument can MOVE to it, and the restriction on MOVE that only elements of a Numeration can be the targets for MERGE. It's this restriction that is responsible for



preventing the object from adjoining to the lower segment of  $vP$  in the step following (159), causing it to surface lower than the underlying position of the subject.

Both the  $\mu$ -variant and this no- $\mu$ -variant are found in the literature. However, the label “ $\mu$ ” is not very common. More popular is the label “AgrO” — for “object Agreement.” This label expresses the hypothesis that there is a structural relationship between the assignment of structural Case and the agreement relation; it hints at an explanation for the relationship expressed in (152). In what follows, I will adopt the no- $\mu$ -variant, as it appears to be now increasingly popular.





# 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* 5:325–366.
- Bach, Emmon. 1984. Some generalizations of categorial grammars. In *Varieties of formal semantics*, ed. Fred Landman and Frank Velman, 1–23. Dordrecht, The Netherlands: Foris Publications.
- Baker, C. L. 1978. *Introduction to generative transformational syntax*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–416.
- Baker, Mark. 1988. *Incorporation*. Chicago, Illinois: University of Chicago Press.
- Baker, Mark C. 2003. *Lexical categories*. Cambridge: Cambridge University Press.
- Baltin, Mark. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13:1–38.
- Barker, Chris. 1995. *Possessive descriptions*. Stanford University: CSLI Publications.
- Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17:347–354.
- Belletti, Adriana. 1982. On the anaphoric status of the reciprocal construction in Italian. *The Linguistic Review* 2:101–138.
- 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* 3:101–126.

- 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.
- Bobaljik, Jonathan David, and Dianne Jonas. 1996. Subject positions and the roles of tp. *Linguistic Inquiry* 27:195–236.
- Bod, Rens, Jennifer Hay, and Stefanie Jannedy. 2003. *Probabilistic linguistics*. Cambridge, Massachusetts: MIT Press.
- Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.
- 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.
- Bošković, Željko. 2001. Floating quantifiers and theta-role assignment. In *North East Linguistic Society*, ed. Minjoo Kim and Uri Strauss, 59–78. Georgetown University: GLSA.
- Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24:591–656.
- Burzio, Luigi. 1986. *Italian syntax*. Reidel Publishers.
- Carrier, Jill, and Janet H. Randall. 1992. The argument structure and syntactic structure of resultatives. *Linguistic Inquiry* 23:173–234.
- 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, 417–454. 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. 1995a. Bare phrase structure. In *Government binding theory and the minimalist program*, ed. Gert Webelhuth, 383–439. Oxford: Oxford University Press.
- Chomsky, Noam. 1995b. *The minimalist program*. Cambridge, Massachusetts: MIT Press.
- Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row, Publishers.
- Chomsky, Noam, and George Miller. 1963. Finitary models of language users. In *Handbook of Mathematical Psychology 2*, ed. Luce, Bush, and Galanter, 419–491. New York, New York: Wiley and Sons.
- Contreras, Heles. 1984. A note on parasitic gaps. *Linguistic Inquiry* 15:698–701.
- Davidson, Donald. 1967. The logical form of action sentences. In *The logic of decision and action*, ed. Nicholas Rescher, 81–95. Pittsburgh: Pittsburgh University Press.
- Déprez, Vivian. 1989. On the typology of syntactic positions and the nature of chains. Doctoral Dissertation, MIT, Cambridge, Massachusetts.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, Massachusetts: MIT Press.
- Diesing, Molly, and Eloise Jelinek. 1995. Distributing arguments. *Natural Language Semantics* 3:123–176.
- Dikken, Marcel den. 1992. *Particles*. Dordrecht, The Netherlands: Holland Institute of Generative Linguistics.

- Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In *Properties, types and meaning: Semantic issues*, ed. Gennaro Chierchia, Barbara H. Partee, and Raymond Turner. Dordrecht, The Netherlands: Kluwer Academic Press.
- Dryer, Matthew S. 1989. Plural words. *Linguistics* 27:865–895.
- 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* 8:546–561.
- Emonds, Joseph E. 1976. *A transformational approach to English syntax*. New York, New York: Academic Press.
- Fillmore, Charles. 1968. The case for case. In *Universals in linguistic theory*, ed. Emmon Bach and Richard Harms, 1–90. New York, New York: Holt, Rinehart and Winston.
- Frazier, Lyn. 1978. On comprehending sentences: Syntactic parsing strategies. Doctoral Dissertation, University of Connecticut, Storrs.
- Gerdts, Donna. 1981. Object and absolutive in Halkomelem Salish. Doctoral Dissertation, University of California, San Diego.
- Gibson, Edward. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68:1–76.
- 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* 21:633–641.
- Godfrey, J., E. Holliman, and J. McDaniel. 1992. SWITCHBOARD: telephone speech corpus for research and development. In *Proceedings of ICASSP 92*, 517–520. San Francisco.
- Green, Georgia M. 1974. *Semantics and syntactic regularity*. Bloomington, Indiana: Indiana University Press.

- 
- Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. Joseph Greenberg. Cambridge, Massachusetts: MIT Press.
- Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10:279–326.
- Grimshaw, Jane. 1990. *Argument structure*. Cambridge, Massachusetts: MIT Press.
- Gropen, Jess, Steven Pinker, Michelle Hollander, Richard Godberg, and Ronald Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language* 65:203–257.
- Gruber, Jeffrey S. 1965. Studies in lexical relations. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Hackl, Martin. 2000. Comparative quantifiers. Doctoral Dissertation, Massachusetts Institute of Technology.
- 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.
- Harris, Zellig. 1946. From morpheme to utterance. *Language* 22:161–183.
- Holmberg, Anders. 1986. Word order and syntactic features. Doctoral Dissertation, Stockholm.
- 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* 21:427–455.
- Jaeggli, Osvaldo. 1982. *Topics in Romance syntax*. Dordrecht Holland: Foris Publications.

- Johnson, Kyle. 1991. Object positions. *Natural Language and Linguistic Theory* 9:577–636.
- Katz, Jerrold J., and Paul Postal. 1964. *An integrated theory of linguistic descriptions*. Cambridge, Massachusetts: MIT Press.
- 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. Principles of particle constructions. In *Grammatical representation*, ed. Jacqueline Guéron, Hans G. Obenauer, and Jean-Yves Pollock, 101–140. Dordrecht, The Netherlands: Foris Publications.
- Kayne, Richard S. 1984c. Unambiguous paths. In *Connectedness and binary branching*, 129–164. Dordrecht: Foris Publications.
- Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: M.I.T. Press.
- Kimball, J. 1973. Seven principles of surface structure parsing in natural language. *Cognition* 2:15–47.
- Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Koster, Jan. 1975. Dutch as an SOV language. *Linguistic Analysis* 1:111–136.
- 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, 109–137. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Larson, Richard. 1988. On the double object construction. *Linguistic Inquiry* 19:335–392.
- 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* 21:103–122.
- 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, Beth, and Malka Rappaport. 1986. The formation of adjectival passives. *Linguistic Inquiry* 17:623–661.
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, Massachusetts: MIT Press.
- Marcus, Mitchell, Beatrice Santorini, and May Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn treebank. In *Computational Linguistics* 19, 313–330.
- Martin, Roger. 2001. Null case and the distribution of PRO. *Linguistic Inquiry* 32:141–166.
- May, Robert. 1985. *Logical form: Its structure and derivation*. Cambridge, Massachusetts: MIT Press.
- Miyagawa, Shigeru. 1989. *Structure and case marking in Japanese*. San Diego: Academic Press.
- 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.
- Partee, Barbara H., and Vladimir Borschev. 2003. Genitives, relational nouns, and argument-modifier ambiguity. In *Modifying adjuncts*, ed. E. Lang, C. Maienborn, and C. Fabricius-Hansen, 67–112. Mouton de Gruyter.
- Perlmutter, David, and Paul Postal. 1984. The 1-advancement exclusiveness law. In *Studies in relational grammar volume 2*, ed. David Perlmutter and Carol Rosen. University of Chicago Press.

- Pesetsky, David. 1982. Paths and categories. Doctoral Dissertation, Massachusetts Institute of Technology.
- Pesetsky, David. 1985. Morphology and logical form. *Linguistic Inquiry* 16:193–248.
- Pesetsky, David. 1995. *Zero syntax*. Cambridge, Massachusetts: MIT Press.
- Pinker, Steven. 1989. *Learnability and cognition. the acquisition of argument structure*. 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.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Massachusetts: MIT Press.
- Roberts, Craige. 1987. Modal subordination, anaphora and distributivity. Doctoral Dissertation, University of Massachusetts, Amherst.
- 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. In *Proceedings of the West Coast Conference on Formal Linguistics*, ed. Jeffrey Goldberg, Susannah MacKay, and Michael T. Wescoat, 273–283. 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, Ken, and Tim Stowell. 1987. Binominal *each*. In *NELS 8*, ed. James Blevins and Julie Carter, 426–450. Toronto: GLSA.
- 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* 29:439–474.
- Sapir, Edward. 1921. *Language: an introduction to the study of speech*. New York: Harcourt, Brace and Co.
- Speas, Margaret. 2000. *Phrase structure in natural language*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19:425–449.
- von Stechow, Arnim. 1996. The different readings of *wieder* “again”: A structural account. *Journal of Semantics* 13:87–138.
- Stowell, Tim. 1978. What was there before there was there. In *Proceedings of the Fourteenth Regional Meeting of the Chicago Linguistic Society*, ed. D. Farkas. Chicago, Illinois: The University of Chicago.
- Stowell, Timothy. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.
- 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. In *The Proceedings of the North Eastern Linguistics Society*, ed. Tim Sherer, 367–382. Universite du Quebec a Montreal: The Graduate Students Linguistic Association, The University of Massachusetts-Amherst.
- Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford: Oxford University Press.
- Wasow, Tom. 1977. Transformations and the lexicon. In *Formal syntax*, ed. Peter Culicover, Tom Wasow, and Adrian Akmajian, 327–360. New York, New York: Academic Press.
- Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11:203–238.
- Zwart, C. Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer Academic Publishers.