Cambridge University Press
978-0-521-10776-1 - Syntactic Chains
Kenneth J. Safir
Excerpt
More information

1 *Grammatical theory and syntactic theory*

How does the human mind work? What is the structure of the human mental faculties? What aspects of these mental skills are innate, and what aspects are acquired? What must humans 'know' before they learn anything, and how does this influence what they are capable of learning? These ambitious questions raise some of the central issues motivating the emergence of cognitive science in the last few decades, although very few fields outside of psychology have hit upon systematic ways to go about investigating these matters. Linguistics, however, is one field that has shown progress in this regard with respect to the somewhat more specific questions that define, at least within the tradition of generative grammar, the central goals of linguistic research: What is the structure of the human language faculty? What must one 'know' in order to learn a language? What does one know when one 'knows' a language? What is the range of possible human languages? These are still ambitious questions, but there is some reason to be optimistic that recent research is making progress. In the last few years, increasingly principled formal theories of syntax have appeared that account for a widening array of disparate data, much of it discovered by asking questions which were unformulable just a few short years ago.

Since the earliest work in generative grammar, a prerequisite for any theory of syntax has been the postulation of a basic set of syntactic relations in terms of which hypotheses about the structure of the innate human language faculty (or, as it is known in the generative tradition, 'Universal Grammar') can be stated. Indeed the postulation of these formal syntactic relations themselves represents a highly significant claim about the 'notation' of Universal Grammar, especially since the rich interdependencies typical of natural language must be expressed in terms of these notations.

This research is, for the most part, an inquiry into one of the newest members of the vocabulary of primitive syntactic relations, the notion 'X is coindexed with Y', and the theory of indexing that has sprung up around it. In particular, I shall construct and motivate a theory of 'syntactic chains' which avoids some of the recent extensions of the relation 'X is coindexed with Y' to independent styles of indexing. If the theory developed here is correct, then, with respect to syntactic chains, syntactic theory need not be enriched, nor its generality weakened, by the introduction of additional primitive indexing relations.

The theory of syntactic chains developed here, moreover, provides us with a program for research that will touch upon many of the central issues in recent syntactic theory, and raise a wide range of empirical issues that will be investigated from a cross-linguistic perspective. It is hoped that many of these investigations are sufficiently detailed that they will remain useful long after some of the theoretical issues that seem important now are replaced by profounder questions.

The business of constructing and defending this theory of chains begins in the next chapter, but first it is necessary to provide a context for this project by being more explicit about what is meant by a 'theory of grammar,' and by reviewing some recent developments in syntactic theory that will figure prominently in the rest of my discussion.

1.1 The theory of grammar

The goal of linguistic theory is to provide an accurate characterization of the innate human language faculty, or, to use the term introduced above, 'Universal Grammar' (UG). Chomsky (1981a) has described the basic problem posed by this goal as follows,

> The theory of UG must meet two obvious conditions. On the one hand, it must be compatible with the diversity of existing (indeed, possible) grammars. At the same time, UG must be sufficiently constrained and restrictive in the options it permits so as to account for the fact that each of these grammars develops in the mind on the basis of quite limited evidence . . . it is a near certainty that fundamental properties of the attained grammars are radically underdetermined by the evidence available to the language learner and must therefore be attributed to UG itself (p. 30).

I shall assume Universal Grammar to be a set of principles that hold of every language ('universal principles of grammar') and a set of yes/no options ('parameters') that break up the classes of possible languages into intersecting sets. The language learner is presumed to have the universal principles of grammar as part of his or her mental equipment at birth, as well as a schema of parameters that have marked and unmarked values. In learning a language 'X,' the language learner must fix the values for all of the parameters of X on the basis of the limited data to which he or she is exposed, and must acquire a lexicon for X (also presumably constrained by the nature of the universal principles of grammar and parameters). The adult form of the linguistic competence of a native speaker of X includes knowledge of the lexicon of X (Lx) and fixed ('yes' or 'no') values for the parameters of X (Px); the rest of the native speaker's knowledge of X should follow from the interaction of Px and Lx with the universal principles of grammar. This interaction is called the 'core grammar of X.'

One highly successful strategy for uncovering properties of UG has been the detailed study and comparison of adult, or 'final state' grammars, that is to say, grammars as they exist in the minds of speakers who have completed the acquisition process. The final state grammar of a given natural language may be thought of, at the relevant level of abstraction, as an exemplified core grammar. Comparative study of final state grammars has grown increasingly fruitful in recent years as detailed theoretically informed work on particular languages has accumulated, and as the increasing sophistication of linguistic theory has permitted more specific cross-linguistic hypotheses to be constructed and tested. Within this framework of research, UG itself has emerged, methodologically speaking, as the abstraction from final state grammars of principles true of all possible final state grammars. Under the idealization proposed in Chomsky (1965), particularly the assumption that language is instantaneously acquired, it follows that the universal principles of grammar of the final state abstraction form a model of the innate (initial state) human language faculty. Parameters fixed for a given value are, from this point of view, formal properties that hold of classes of final state (again read 'core') grammars.[1]

Part of the focus of this study, particularly Chapter 6, will be to examine the formal properties of some of the parameters that distinguish the Romance languages both from each other, and from the Germanic languages.

The central theme of my research, however, concerns properties of Universal Grammar, namely, the theory of 'syntactic chains,' and the theory of syntactic relations, especially the theory of indexing, within which chains are defined.

1.2 The vocabulary of syntactic relations

The primitive vocabulary of syntactic relations, some of which date back to structuralist grammars, may be stated quite informally as in (1).

(1) Primitive syntactic relations
a. X is (string) adjacent to Y
b. X is in configuration with Y (e.g., sister, daughter)
c. X shares the feature [+F] with Y
d. X is coindexed with Y

Every syntactic relation or interdependency is expressed in terms of one or more of these primitive formal relations.

For example, adjacency is held to be crucial for the statement of rules of contraction, such as those that derive the examples in (2b) and (3b), but not (2a) and (3a).

(2) a. *I probably'm sick
b. I'm probably sick

(3) a. *I wan' John'na leave
b. I wanna leave

Though these issues are more complex than they seem at first (cf. 2.4.1 for discussion), and though appeal to a more abstract notion of adjacency than string adjacency is required, the role of adjacency in these matters is uncontroversial.

The feature-sharing relation is commonly appealed to in attempting to explain the parallel behavior of syntactic constituents on the basis of the fact that they are of the same type. Any treatment of nouns or of noun phrases assumes that these elements are identifiable due to the categorial features they bear. In most recent accounts, Chomsky's (1970) system of syntactic categorial features is assumed.

(4)

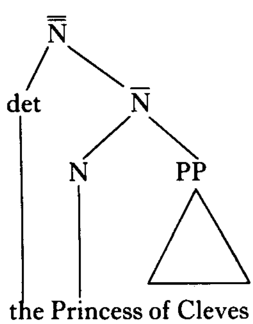
	N	V
noun	+	−
preposition	−	−
adjective	+	+
verb	−	+

Behavior common to verbs and prepositions, for example, includes the fact that they assign Casemarking, whereas adjectives and nouns are generally assumed not to assign Case.[2] Adjectives and verbs, however, more often act as predicates than nouns and prepositions, and never act as 'syntactic arguments' (in a sense to be made more precise in Chapter 2). I shall assume the feature system in (4), but cf. van Riemsdijk (1981), Jackendoff (1977) and Stowell (1981) for further discussion and references.

Configurational relations have played a particularly prominent role in recent theoretical developments, especially in formulations of the '$\bar{X}$ system,' 'government' and 'C-command.'

The basic idea of $\bar{X}$ theory, introduced in Chomsky (1970) and further developed in the references cited above, is that the rewriting relation 'X rewrites as Y' (X → Y) is constrained as to the possible values of X and Y. For a given lexical category, $[\mp N, \mp V]$, every node dominating $[\mp N, \mp V]$ up through a certain number of dominating nodes must bear the same categorial features. Thus in the diagram below of an NP, it is assumed that there are two 'projections,' $\bar{N}$ and $\bar{\bar{N}}$, of the 'head' N (where 'N' now stands for 'noun,' and not the syntactic feature 'N' which appears on the chart in (4)).

(5)



The highest projection of a head 'X' is the 'maximal projection of X,' where 'X' may be taken to be any lexical head such as N, P(reposition), A(djective), or V(erb). These sorts of relations are expressed in terms of a combination of configurational and feature-sharing relations. Further extension of this sort of relation might be to define 'object of' as being, say, 'NP sister of X (where X is a head)' as opposed to, say, 'NP sister of X^{max}.' In later chapters I will discuss some further configurational definitions of this nature.

One configurational relation that has been at the center of many new theoretical developments is 'government.' Government interacts with many 'subtheories' of grammar (cf. 1.4), including binding theory, Case theory, and the set of assumptions surrounding the Empty Category Principle, all of which will be discussed at some length in the course of my presentation. In order to understand how this relation is motivated, however, it is necessary to have some sense of its formal character.

The basic idea or 'core notion' of government, as Chomsky (1981a) has put it, is essentially the traditional idea that a head is in a special relation with its complements. Chomsky adds, however, that 'the operative notion involves structural configurations generalizing the core notion' (p. 163). One such generalization of the traditional notion results from treating the notion of 'head' in the sense of $\bar{X}$ theory. Moreover, if the $\bar{X}$ system is extended to treat INFL(ection) as the head of $\bar{S}$,[3] then, within the context of $\bar{X}$ theory, the core notion is already extended beyond the traditional sense. (The class of $\bar{X}$ theory heads thus includes INFL, N, A, V and P.) The core notion is extended even further, again within the context of $\bar{X}$ theory, by the definition of government I shall be assuming in this study, which is essentially that of Aoun and Sportiche (1981). (For some alternative definitions of government, cf. Rouveret and Vergnaud (1980), Jaeggli (1980b), Chomsky (1980), Kayne (1981a) and Baltin (1982).)

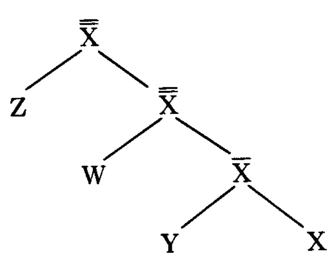
(6) Government
α governs γ in a structure [$\beta \ldots \gamma \ldots \alpha \ldots \gamma \ldots$], where,
i. $\alpha = X^{\circ}$
ii. Where ϕ is a maximal projection, ϕ dominates α if and only if ϕ dominates γ.

I depart from the Aoun and Sportiche definition in my interpretation of 'maximal projection,' however. Consider the definitions in (7) and (8).

(7) Base maximal projection[4]
X^{n} is the base maximal projection of X° if n is the highest value for the category X in the base.

(8) Maximal projection
A maximal projection of X is the highest projection of X°, X^{n}, where X^{n} is base maximal.

(9)

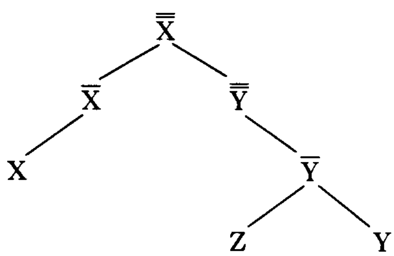


The notion assumed by Aoun and Sportiche corresponds to 'base maximal projection,' and thus X governs Y and W in their account, but X does not govern Z because X and Z are not both dominated by all of the same base maximal projections of X (i.e., there are two base maximal projections of X, and Z is only dominated by one of them). Under the interpretation adopted here, the 'maximal projection' relevant to the definition of government is uniquely the node that dominates Z, hence X governs Z.[5] Given either interpretation of the notion 'maximal projection,' however, it follows that

(10) Maximal projections are absolute barriers to government.

Thus in diagram (11), X does not govern any daughter of $\bar{Y}$ if $\bar{Y}$ is a maximal projection.

(11)



The basic idea incorporated in this definition is that a head governs all of its complements within the domain of its own maximal projection, but does not govern those within the domain of any other maximal projection.

One empirical consequence of government as it is defined in (10) may be observed with respect to Case assignment, if Case assignment is to be constrained by government, as has been proposed by Rouveret and Vergnaud (1980) and Chomsky (1980). Following these authors, a verb which assigns Accusative Case must govern an NP in order to assign Case to that NP successfully. In this light, consider (12).

(12) a. John believed him/*he
b. John believed $[_{\bar{S}}$ he/*him was innocent]
c. $[_{\bar{S}}$ For $[_{S}$ him to leave]] was foolish

Under the assumption that $\bar{S}$ is the maximal projection of INFL (cf. n. 3), *believe* does not govern the pronoun in (12b), since a maximal projection intervenes between *believe* and the pronoun. In (12c), however, the prepositional complementizer *for* is within $\bar{S}$, and can govern the subject of the infinitive, just as a verb can govern its object as in (12a) (but cf. n. 7).

One more configurational notion which will play a role in my discussion is that of 'C-command.' I adopt a formulation of C-command very similar to that of Aoun and Sportiche (1981), which I define as in (13).[6]

(13) C-command
α C-commands β if the first maximal projection dominating α also dominates β, and α does not contain β.

A typical example of the operation of C-command concerns the contrast in (14).

(14) a. **he* likes the woman [$_{\bar{S}}$ who kissed *John*]
b. [$_{\bar{S}}$ the woman [$_{\bar{S}}$ who kissed *John*] [$_{VP}$ liked *him*]]

It is well known that a name cannot be coreferent with a pronoun that C-commands it. In (14b), the first maximal projection dominating *John* is the $\bar{S}$ within the subject relative clause, while in (14a), the first maximal projection dominating *he* is the matrix $\bar{S}$, and the matrix $\bar{S}$, of course, dominates everything in the sentence. Thus *he* C-commands the name *John* in (14a), and *he* and *John* must be disjoint in reference, whereas *him* and *John* can corefer in (14b) because neither NP C-commands the other (cf. Lasnik (1976) and Reinhart (1976)). C-command is also considered to be a crucial factor for determining quantifier scope, a matter which will be touched on in Chapters 2 and 5.

The variety of relations definable on the primitive syntactic relations in (1) is already vast, and permits a great deal of descriptive precision. Any addition to the class of primitive syntactic relations is therefore to be avoided, since it increases the class of possible syntactic relations that can be expressed (that is to say, the explanatory force of the relations in (1) is weakened). The primitive relation 'X is coindexed with Y' will be viewed from this perspective in the next chapter, where syntactic relations that depend crucially on coindexing are defined.

1.3 From systems of rules to systems of principles

As Chomsky (1981a) has pointed out, the recent shift in focus from systems of rules to systems of principles is perhaps the most striking and most promising theoretical development of the last decade. There are a number of new directions of investigation resulting from this shift, some of which enhance the explanatory role of syntactic relations in ways relevant to the analysis of indexing in the next chapter.

The shift from rules to principles had its origin, in part, in the formulation of general constraints on transformations, such as those in Ross (1967). Nonetheless, most investigation that followed still aimed at the discovery of rules characterizing constructions and generalizations across the latter, such as the Complex Noun Phrase Constraint. The unification of some of these generalizations under more abstract principles, such as 'subjacency' and 'opacity' (the Tensed S Condition and the Specified Subject Condition), in Chomsky (1973) marked a change of focus towards abstract theorems and their empirical consequences, rather than, or in addition to, the more data-driven sorts of generalizations across descriptive rules that characterized much of earlier research.

Another part of this shift, however, was the abandonment of the hypothesis that all semantic interpretation is at D-structure. Under the latter hypothesis (Katz and Postal (1964)), it had to be assumed that transformational rules could not apply so as to produce the wrong output at surface structure. Many of the complexities of transformational rules and phrase structure rules were then justified as a means of avoiding the generation of ungrammatical strings after semantic interpretation. The shift to surface interpretation in the early seventies (cf. Jackendoff (1972) and Chomsky (1972)) made it possible to marshal general interpretive constraints to rule out overgeneration, that is to say, the idea emerged that the constraints of one component might filter out the overabundant production of other components.

The first casualty of this shift were complex transformational rules, particularly in terms of their structural descriptions (which limited the contexts where they could apply), their obligatory or non-obligatory character, and their orderings with respect to one another. Chomsky's (1976) introduction of 'minimal factorization,' and finally the even simpler 'Move α' in Chomsky (1980), places the burden of explanation on general principles rather than on specific rules encoding constructions.

It should be noted, however, that although Move α obliterated ordered transformational rules, it did not vitiate (indeed it enhanced) the claim that derivations have beginnings and ends and pass through distinct levels at which major principles hold. Thus the ordering of rules largely gave way to the ordering of components and levels, and to the means by which one level is mapped onto another. A very simplified version of the Chomsky and Lasnik (1977) model is presented below.

(15)

D-structure
↓
Move α
↓
S-structure

(Phonetic Form) PF	LF (Logical Form)
Filters	Opacity
Surface structure	Quantifier interpretation

An important aspect of the overgeneration/filtering approach that emerged in Chomsky and Lasnik's work is the role of conspiracies in ruling out the various ungrammatical members of paradigms generable by base rules and Move α. While many of the particular filters and results achieved by conspiracies between filters that they proposed have since been derived by more general principles (e.g., the *that-e* Filter, see Chapter 2), the conspiratorial force of interacting syntactic principles has since become a major theoretical focus.

Notice, however, that emerging also in this shift from rules to principles, and out from behind the secondary syntactic relations such as government and C-command, is a new explanatory role for the basic vocabulary of syntactic relations. For example, the primitive relation of configuration is now, in effect, conditioned by $\bar{X}$ theory, thus allowing a vastly decreased class of possible phrase structures, yet dominance and sisterhood remain central relations increasingly associated with diagnostic properties precisely because of the impoverished descriptive power of base rules. In some recent studies, base rules have been virtually eliminated altogether in favor of the interaction of other principles and components, thereby increasing the explanatory force of configurationality.

Although it is beyond the scope of this study, this discussion could be extended considerably with respect to languages for which it has been claimed that configurationality is not a relevant primitive relation, i.e., the suggestion is that configurationality is parameterized (Hale (1978)). Although I am sceptical about the claim that 'non-configurational languages' exist (though it seems plausible that minimally configurational languages do, cf. Hale (1983), the fact that such a parameter could be proposed exemplifies the new explanatory role of the primitive syntactic relations as they have become more closely identified with diagnostic properties.