

## Chapter 1

INTRODUCTION

1.0. The past decade of research on transformational grammar has substantiated amply, to my mind, the claim that the optimal framework for the description of syntactic facts is a set of rules, of two types: context-free phrase structure rules, which generate an infinite set of highly abstract formal objects, underlying (or deep) phrase markers; and grammatical transformations, which map underlying phrase markers onto an infinite set of objects of roughly the same formal character, superficial (or surface) phrase markers.<sup>1</sup> Within this framework, an evaluation measure is provided which must select, from a set of observationally adequate grammars of some language -- i.e., grammars which all generate the observed set of grammatical sentences of the language -- the descriptively adequate grammar -- the grammar which makes correct predictions about strings of words not yet observed, and can thus be said to reflect linguistic knowledge of speakers of the language.<sup>2</sup> Such knowledge includes intuitions about the immediate constituents of sentences, about similarity among constituents, and about relatedness between sentences. For instance, a descriptively adequate grammar of English would have to predict the following facts about sentence (1.1):

(1.1) A gun which I had cleaned went off.

- a) The main constituent break occurs between cleaned and went; I is a constituent; which I is not; etc.

- b) The constituent a gun which I had cleaned is a constituent of the same kind as the constituent I. Similarly, went off is the same type of constituent as had cleaned, and neither is of the same type as I, a, or off.
- c) Sentence (1.1) is related to sentence (1.2).

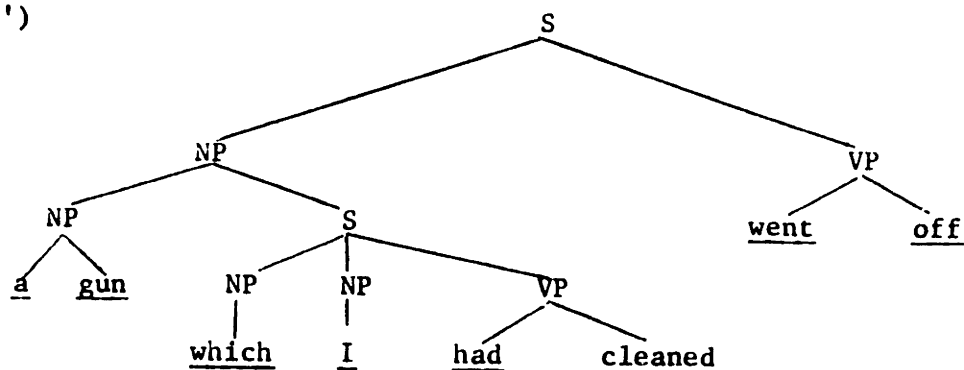
(1.2) A gun went off which I had cleaned.

Within a transformational grammar, intuitions of relatedness between sentences are reconstructed by deriving sets of related sentences from the same or highly similar underlying phrase markers by means of slightly differing sets of transformations. As a first approximation, we could postulate a rule like (1.3) to convert the structure underlying (1.1) to the one underlying (1.2)<sup>3</sup> (here and elsewhere I will give rules and tree diagrams in a simplified form, as long as it makes no difference for the point under discussion):

$$\begin{array}{rcccc}
 (1.3) & [NP & S] & VP & \\
 & & NP & & OPT \\
 & 1 & 2 & 3 & \implies \\
 & 1 & 0 & 3 + 2 &
 \end{array}$$

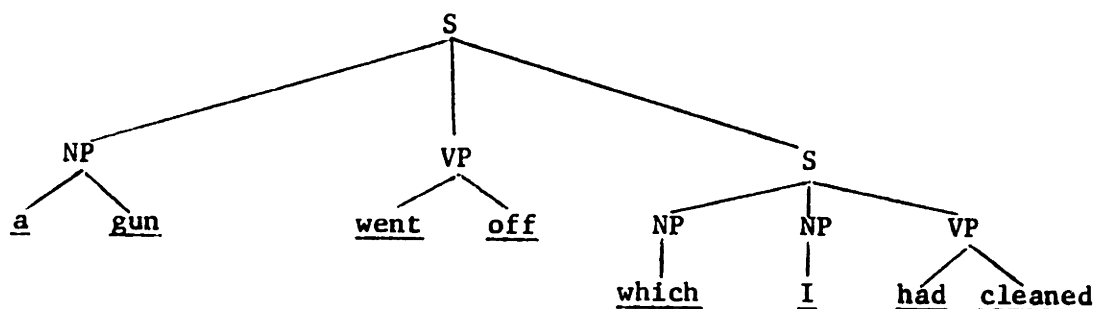
where the phrase marker (P-Marker) associated with (1.1) can be represented as a tree diagram of roughly the following form<sup>4</sup>:

(1.1')



Rule (1.3) would convert (1.1') into the derived P-Marker (1.2')

(1.2')



It is fairly easy to demonstrate that the present evaluation measure gives a higher rating to a grammar which has (1.1') as an underlying P-Marker and derives (1.2') from it by using (1.3), than to one which assumes (1.2') is basic; but I will not undertake such a demonstration here, since the point at issue is more general, and these rules I propose are only supposed to illustrate it, not to constitute a complete analysis.

Now consider the sentences (1.4) and (1.5).

(1.4) I gave a gun which I had cleaned to my brother.

(1.5) I gave a gun to my brother which I had cleaned.

To relate (1.4) and (1.5) -- again, I omit the argument which would prove that (1.5) must derive from (1.4) -- some rule like (1.6) would be necessary.

$$\begin{array}{rccccccc}
 (1.6) & \text{NP} & \text{V} & \text{[NP - S]} & - & \text{PP} & \\
 & \underbrace{\hspace{1.5cm}} & & \text{NP} & & & \text{OPT} \\
 & 1 & & 2 & & 3 & \implies \\
 & 1 & & 0 & & 3 + 2 & 
 \end{array}$$

By the provisions of the evaluation measure, we are forced to collapse rules which are similar in certain ways, and (1.3) and (1.6) collapse to yield (1.7):

$$\begin{array}{rccccccc}
 (1.7) & \left\{ \begin{array}{c} \emptyset \\ \text{NP} \quad \text{V} \end{array} \right\} & - & \text{[NP - S]} & - & \left\{ \begin{array}{c} \text{VP} \\ \text{PP} \end{array} \right\} & \\
 & \underbrace{\hspace{1.5cm}} & & \text{NP} & & & \text{OPT} \\
 & 1 & & 2 & & 3 & \implies \\
 & 1 & & 0 & & 3 + 2 & 
 \end{array}$$

Consideration of sentences like (1.8) and (1.9).

(1.8) He let the cats which were meowing out.

(1.9) He let the cats out which were meowing.

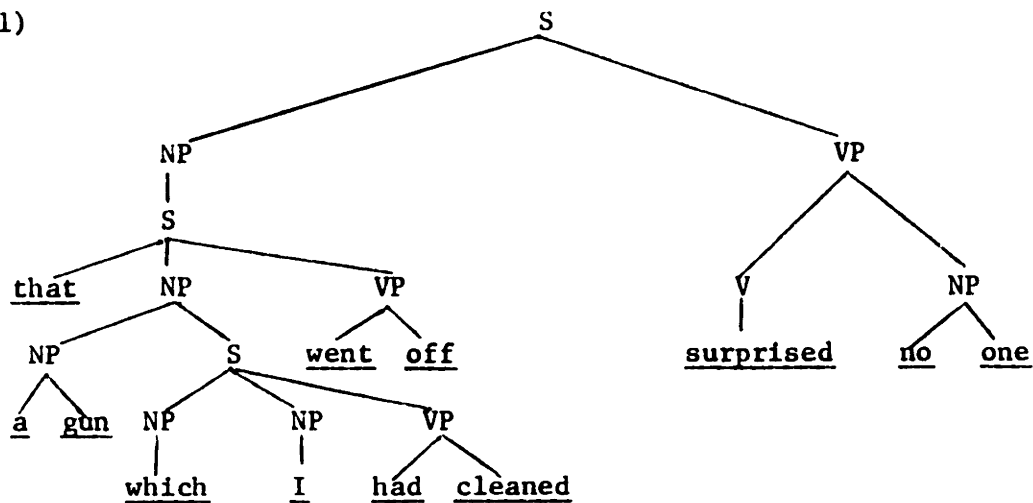
and similar sentences might lead one to reformulate (1.7) as an even more general rule, (1.10), which I will call Extrapolation from NP:

(1.10) Extrapolation from NP

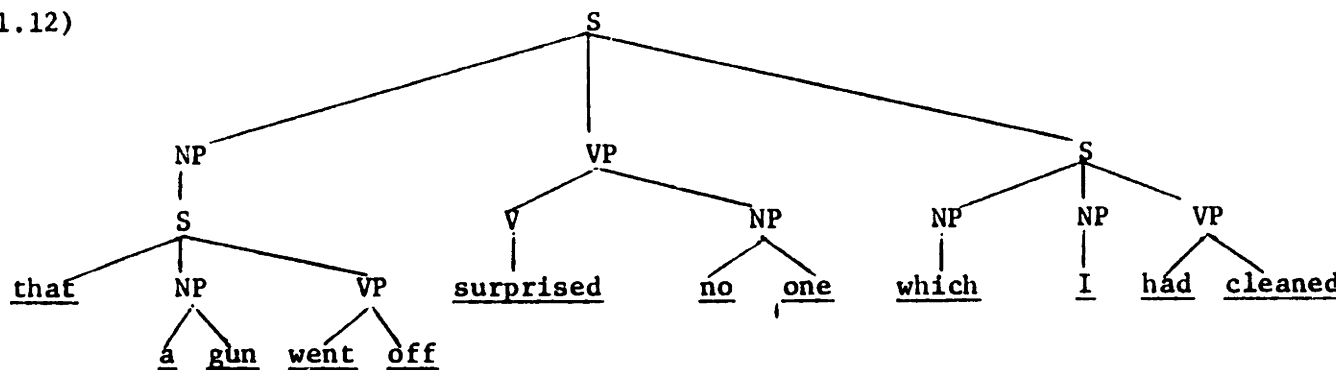
$$\begin{array}{rccccccc}
 & \text{X} & \text{[NP - S]} & - & \text{Y} & & \text{OPT} \\
 & & \text{NP} & & \text{NP} & & \\
 & \underbrace{\hspace{1.5cm}} & & & & & \implies \\
 & 1 & 2 & & 3 & & \\
 & 1 & 0 & & 3 + 2 & & 
 \end{array}$$

The symbols X and Y in (1.10) are variables which range over all strings, including the null string. With them, the rule as it stands is much too powerful. For instance, (1.10) would convert (1.11) into the ungrammatical (1.12).

(1.11)



(1.12)



The fact is that an extraposed clause may never be moved outside "the first sentence up," in the obvious interpretation of this phrase, and there are a number of ways of incorporating this fact into



(1.14) What did Bill buy?

What did you force Bill to buy?

What did Harry say you had forced Bill to buy?

What was it obvious that Harry said you had  
forced Bill to buy?

A moment's reflection should convince anyone that it is impossible to replace the variable X in (1.13) by some such disjunction as that contained in (1.7): rule (1.13) is not stateable without variables. And yet, just as was the case with rule (1.10), Extrapolation from NP, it is easy to see that (1.13) is far too strong, for it will generate infinitely many non-sentences, such as those in (1.15).

(1.15) \* What did Bill buy potatoes and?

\* What did that Bill wore surprise everyone?

\* What did John fall asleep and Bill wear?

1.1. Sentences and non-sentences like those in (1.14) and (1.15) show that some rules must contain variables but that somehow the power of these variables must be restricted. It is the purpose of this thesis to try to justify a set of constraints on variables, which I will propose in detail in subsequent chapters. There are doubtless many constraints on variables which are peculiar to individual languages, and possibly some which are even peculiar to some rule in some particular language, but I have by and large avoided detailed discussion of these and have instead concentrated my research on constraints which I suspect to be universal.

It is obvious that the limited character of presently available syntactic knowledge reduces drastically the chances of survival of any universals which can be formulated today, for the study of syntax is truly in its infancy. But it will be seen below that the constraints on variables which I will propose are often of such a complex nature that to state them as constraints on rules in particular languages would greatly increase the power of transformational rules and of the kinds of operations on P-Markers they could perform. But to assume more powerful apparatus in a theory than can be shown to be necessary is contrary to basic tenets of the philosophy of science, and so I will tentatively assume that many of the constraints I have arrived at in my investigations of the few languages I am familiar with are universal. It is easy to prove me mistaken in this assumption: if languages can be found whose rules are not subject to these constraints, then the apparatus in theory of generative grammar which provides for the description of language - particular facts will have to be strengthened so that rules like the question transformation in English, (1.13), for instance, can be stated and correctly restricted to exclude ungrammatical sentences like those in (1.15). But until such disconfirming evidence arises, the assumption of a weaker theory for particular languages is dictated by principles of the philosophy of science.

It is probably unnecessary to point out that it is commonplace to limit the power of the apparatus which is available for the description of particular languages by "factoring out" of individual



grammars, principles, conditions, conventions and concepts which are necessary in all grammars: to factor out in this manner is to construct a theory of language. So, for example, when the principle of operation of the syntactic transformational cycle has been specified in linguistic theory, it is unnecessary to include another description of this principle in a grammar of French. And so it is also with such well-known notions as free variation, grammatical sentence, constituent, coordinate structure, verb, and many others. The present work should be looked upon as an attempt to add to this list a precise specification of the notion syntactic variable. This notion is crucial for the theory of syntax, for without it the most striking fact about syntactic processes - the fact that they may operate over indefinitely large domains - cannot be captured. And since almost all transformations either are most generally stated, or can only be stated, with the help of variables, no transformation which contains variables in its structural index will work properly until syntactic theory has provided variables which are neither too powerful nor too weak. It is easy to construct counterexamples such as those in (1.15) for almost every transformation containing variables that has ever been proposed in the literature on generative grammar.<sup>1</sup> It is for this reason that attempts to constrain variables, like those which will be discussed in Chapters 2, 4, and 5, are so important: without the correct set of constraints, it is impossible to formulate almost all syntactic rules precisely, unless one is willing to so greatly

increase the power of the descriptive apparatus that every variable in every rule can be constrained individually. But one pursuing this latter course will soon come to realize that many of the constraints he imposes on individual variables must be stated again and again; that he is missing clear generalizations about language. Thus, the latter course must be abandoned: the only possible course is to search for universal constraints. This thesis is devoted to that search.

1.2. The outline of this work is as follows. In Chapter 2, I will discuss the only previous attempts to limit the power of variables which I know of<sup>5</sup>, Chomsky's A-over-A principle, and two conditions subsequently proposed by him, and demonstrate that they are too strong in some respects and too weak in others. In Chapter 3, I will discuss a notion which will prove indispensable in stating the universal constraints: the notion of node deletion, or tree pruning. In Chapter 4, I state and discuss two putatively universal constraints on variables, which overcome the inadequacies in the principles discussed in Chapter 2, and several less general constraints. The notion of bounding is introduced in Chapter 5. In Chapter 6, I discuss briefly a number of rules and show that these rules are<sup>1</sup> subject to the constraints of Chapter 4, but that not all transformations are subject to these constraints. The question is discussed as to what formal features of rules determine whether the variables in them are subject to the constraints or not. Chapter 7 is a brief recapitulation of the results of the thesis.

## Chapter 1

FOOTNOTES

1. For an excellent introductory article on the difference between underlying and superficial structure, cf. Postal (1964). A more technical and far more complete exposition is given in Chomsky (1965).
2. For further discussion of the notions of observational and descriptive adequacy, cf. Chomsky (1964b).
3. My notation for transformations follows that of Rosenbaum (1965), except where otherwise noted.
4. The assumption that relative clauses are introduced in the deep structure by the rule  $NP \rightarrow NP S$  will be justified in Lakoff and Ross (in preparation b).
5. Except Langacker's notion of command (Langacker (1966)) and Klima's notion in construction with (Klima (1964)), which will be discussed separately in §5 below, in connection with the notion of bounding.