TABLE OF CONTENTS


Introduction to Second Edition by David W. Lightfoot . . . v
Preface . . . . . . . . . . . . . . . . . . . . . . 5
1. Introduction . . . . . . . . . . . . . . . . . . . 11
2. The Independence of Grammar . . . . . . . . . . . 13
3. An Elementary Linguistic Theory . . . . . . . . . . . 18
4. Phrase Structure . . . . . . . . . . . . . . . . . . 26
5. Limitations of Phrase Structure Description . . . . . . 34
6. On the Goals of Linguistic Theory . . . . . . . . . . 49
7. Some Transformations in English . . . . . . . . . . . 61
8. The Explanatory Power of Linguistic Theory . . . . . . 85
9. Syntax and Semantics . . . . . . . . . . . . . . . 92
10. Summary . . . . . . . . . . . . . . . . . . . . . 106
11. Appendix I: Notations and Terminology . . . . . . . . 109
12. Appendix II: Examples of English Phrase Structure and Transformational Rules . . . . . . . . . . . . . . . . 111
Bibliography . . . . . . . . . . . . . . . . . . . . . 115

# 1

# INTRODUCTION

Syntax is the study of the principles and processes by which sentences are constructed in particular languages. Syntactic investigation of a given language has as its goal the construction of a grammar that can be viewed as a device of some sort for producing the sentences of the language under analysis. More generally, linguists must be concerned with the problem of determining the fundamental underlying properties of successful grammars. The ultimate outcome of these investigations should be a theory of linguistic structure in which the descriptive devices utilized in particular grammars are presented and studied abstractly, with no specific reference to particular languages. One function of this theory is to provide a general method for selecting a grammar for each language, given a corpus of sentences of this language.

The central notion in linguistic theory is that of "linguistic level." A linguistic level, such as phonemics, morphology, phrase structure, is essentially a set of descriptive devices that are made available for the construction of grammars; it constitutes a certain method for representing utterances. We can determine the adequacy of a linguistic theory by developing rigorously and precisely the form of grammar corresponding to the set of levels contained within this theory, and then investigating the possibility of constructing simple and revealing grammars of this form for natural languages. We shall study several different conceptions of linguistic structure in this manner, considering a succession of linguistic levels of increasing complexity which correspond to more and more powerful modes of grammatical description; and we shall attempt to show that linguistic theory must contain at least these levels if it is to provide, in particular, a satisfactory grammar of English. Finally, we shall suggest that this purely formal investigation of the structure of language has certain interesting implications for semantic studies.[1]

[1] The motivation for the particular orientation of the research reported here is discussed below in § 6.

2

# THE INDEPENDENCE OF GRAMMAR

**2.1** From now on I will consider a *language* to be a set (finite or infinite) of sentences, each finite in length and constructed out of a finite set of elements. All natural languages in their spoken or written form are languages in this sense, since each natural language has a finite number of phonemes (or letters in its alphabet) and each sentence is representable as a finite sequence of these phonemes (or letters), though there are infinitely many sentences. Similarly, the set of 'sentences' of some formalized system of mathematics can be considered a language. The fundamental aim in the linguistic analysis of a language L is to separate the *grammatical* sequences which are the sentences of L from the *ungrammatical* sequences which are not sentences of L and to study the structure of the grammatical sequences. The grammar of L will thus be a device that generates all of the grammatical sequences of L and none of the ungrammatical ones. One way to test the adequacy of a grammar proposed for L is to determine whether or not the sequences that it generates are actually grammatical, i.e., acceptable to a native speaker, etc. We can take certain steps towards providing a behavioral criterion for grammaticalness so that this test of adequacy can be carried out. For the purposes of this discussion, however, suppose that we assume intuitive knowledge of the grammatical sentences of English and ask what sort of grammar will be able to do the job of producing these in some effective and illuminating way. We thus face a familiar task of explication of some intuitive concept — in this case, the concept "grammatical in English," and more generally, the concept "grammatical."

Notice that in order to set the aims of grammar significantly it is sufficient to assume a partial knowledge of sentences and non-sentences. That is, we may assume for this discussion that certain sequences of phonemes are definitely sentences, and that certain other sequences are definitely non-sentences. In many intermediate cases we shall be prepared to let the grammar itself decide, when the grammar is set up in the simplest way so that it includes the clear sentences and excludes the clear non-sentences. This is a familiar feature of explication.[1] A certain number of clear cases, then, will provide us with a criterion of adequacy for any particular grammar. For a single language, taken in isolation, this provides only a weak test of adequacy, since many different grammars may handle the clear cases properly. This can be generalized to a very strong condition, however, if we insist that the clear cases be handled properly for *each* language by grammars all of which are constructed by the same method. That is, each grammar is related to the corpus of sentences in the language it describes in a way fixed in advance for all grammars by a given linguistic theory. We then have a very strong test of adequacy for a linguistic theory that attemps to give a general explanation for the notion "grammatical sentence" in terms of "observed sentence," and for the set of grammars constructed in accordance with such a theory. It is furthermore a reasonable requirement, since we are interested not only in particular languages, but also in the general nature of Language. There is a great deal more that can be said about this crucial topic, but this would take us too far afield. Cf. § 6.

**2.2** On what basis do we actually go about separating grammatical sequences from ungrammatical sequences? I shall not attempt to

[1] Cf., for example, N. Goodman, *The structure of appearance* (Cambridge, 1951), pp. 5–6. Notice that to meet the aims of grammar, given a linguistic theory, it is sufficient to have a partial knowledge of the sentences (i.e., a corpus) of the language, since a linguistic theory will state the relation between the set of observed sentences and the set of grammatical sentences; i.e., it will define "grammatical sentence" in terms of "observed sentence," certain properties of the observed sentences, and certain properties of grammars. To use Quine's formulation, a linguistic theory will give a general explanation for what 'could' be in language on the basis of "what *is* plus *simplicity* of the laws whereby we describe and extrapolate what is". (W. V. Quine, *From a logical point of view* [Cambridge, 1953], p. 54). Cf. § 6.1.



give a complete answer to this question here (cf. §§ 6.7), but I would like to point out that several answers that immediately suggest themselves could not be correct. First, it is obvious that the set of grammatical sentences cannot be identified with any particular corpus of utterances obtained by the linguist in his field work. Any grammar of a language will *project* the finite and somewhat accidental corpus of observed utterances to a set (presumably infinite) of grammatical utterances. In this respect, a grammar mirrors the behavior of the speaker who, on the basis of a finite and accidental experience with language, can produce or understand an indefinite number of new sentences. Indeed, any explication of the notion "grammatical in L" (i.e., any characterization of "grammatical in L" in terms of "observed utterance of L") can be thought of as offering an explanation for this fundamental aspect of linguistic behavior.

**2.3** Second, the notion "grammatical" cannot be identified with "meaningful" or "significant" in any semantic sense. Sentences (1) and (2) are equally nonsensical, but any speaker of English will recognize that only the former is grammatical.

(1) Colorless green ideas sleep furiously.
(2) Furiously sleep ideas green colorless.

Similarly, there is no semantic reason to prefer (3) to (5) or (4) to (6), but only (3) and (4) are grammatical sentences of English.

(3) have you a book on modern music?
(4) the book seems interesting.
(5) read you a book on modern music?
(6) the child seems sleeping.

Such examples suggest that any search for a semantically based definition of "grammaticalness" will be futile. We shall see, in fact, in § 7, that there are deep structural reasons for distinguishing (3) and (4) from (5) and (6); but before we are able to find an explanation for such facts as these we shall have to carry the theory of syntactic structure a good deal beyond its familiar limits.

**2.4** Third, the notion "grammatical in English" cannot be identified in any way with the notion "high order of statistical approximation to English." It is fair to assume that neither sentence (1) nor (2) (nor indeed any part of these sentences) has ever occurred in an English discourse. Hence, in any statistical model for grammaticalness, these sentences will be ruled out on identical grounds as equally 'remote' from English. Yet (1), though nonsensical, is grammatical, while (2) is not. Presented with these sentences, a speaker of English will read (1) with a normal sentence intonation, but he will read (2) with a falling intonation on each word; in fact, with just the intonation pattern given to any sequence of unrelated words. He treats each word in (2) as a separate phrase. Similarly, he will be able to recall (1) much more easily than (2), to learn it much more quickly, etc. Yet he may never have heard or seen any pair of words from these sentences joined in actual discourse. To choose another example, in the context "I saw a fragile—," the words "whale" and "of" may have equal (i.e., zero) frequency in the past linguistic experience of a speaker who will immediately recognize that one of these substitutions, but not the other, gives a grammatical sentence. We cannot, of course, appeal to the fact that sentences such as (1) 'might' be uttered in some sufficiently far-fetched context, while (2) would never be, since the basis for this differentiation between (1) and (2) is precisely what we are interested in determining.

Evidently, one's ability to produce and recognize grammatical utterances is not based on notions of statistical approximation and the like. The custom of calling grammatical sentences those that "can occur", or those that are "possible", has been responsible for some confusion here. It is natural to understand "possible" as meaning "highly probable" and to assume that the linguist's sharp distinction between grammatical and ungrammatical[2] is motivated by a feeling that since the 'reality' of language is too complex to be described completely, he must content himself with a schematized

[2] Below we shall suggest that this sharp distinction may be modified in favor of a notion of levels of grammaticalness. But this has no bearing on the point at issue here. Thus (1) and (2) will be at different levels of grammaticalness even if (1) is assigned a lower degree of grammaticalness than, say, (3) and (4); but they will be at the same level of statistical remoteness from English. The same is true of an indefinite number of similar pairs.



version replacing "zero probability, and all extremely low probabilities, by *impossible*, and all higher probabilities by *possible*."[3] We see, however, that this idea is quite incorrect, and that a structural analysis cannot be understood as a schematic summary developed by sharpening the blurred edges in the full statistical picture. If we rank the sequences of a given length in order of statistical approximation to English, we will find both grammatical and ungrammatical sequences scattered throughout the list; there appears to be no particular relation between order of approximation and grammaticalness. Despite the undeniable interest and importance of semantic and statistical studies of language, they appear to have no direct relevance to the problem of determining or characterizing the set of grammatical utterances. I think that we are forced to conclude that grammar is autonomous and independent of meaning, and that probabilistic models give no particular insight into some of the basic problems of syntactic structure.[4]

[3] C. F. Hockett, *A manual of phonology* (Baltimore, 1955), p. 10.

[4] We return to the question of the relation between semantics and syntax in §§ 8, 9, where we argue that this relation can only be studied after the syntactic structure has been determined on independent grounds. I think that much the same thing is true of the relation between syntactic and statistical studies of language. Given the grammar of a language, one can study the use of the language statistically in various ways; and the development of probabilistic models for the use of language (as distinct from the syntactic structure of language) can be quite rewarding. Cf. B. Mandelbrot, "Structure formelle des textes et communication: deux études," *Word* 10.1–27 (1954); H. A. Simon, "On a class of skew distribution functions," *Biometrika* 42.425–40 (1955).

One might seek to develop a more elaborate relation between statistical and syntactic structure than the simple order of approximation model we have rejected. I would certainly not care to argue that any such relation is unthinkable, but I know of no suggestion to this effect that does not have obvious flaws. Notice, in particular, that for any $n$, we can find a string whose first $n$ words may occur as the beginning of a grammatical sentence $S_1$ and whose last $n$ words may occur as the ending of some grammatical sentence $S_2$, but where $S_1$ must be distinct from $S_2$. For example, consider the sequences of the form "the man who ... are here," where ... may be a verb phrase of arbitrary length. Notice also that we can have new but perfectly grammatical sequences of word classes, e.g., a sequence of adjectives longer than any ever before produced in the context "I saw a — house." Various attempts to explain the grammatical-ungrammatical distinction, as in the case of (1), (2), on the basis of frequency of sentence type, order of approximation of word class sequences, etc., will run afoul of numerous facts like these.

3

# AN ELEMENTARY LINGUISTIC THEORY

**3.1** Assuming the set of grammatical sentences of English to be given, we now ask what sort of device can produce this set (equivalently, what sort of theory gives an adequate account of the structure of this set of utterances). We can think of each sentence of this set as a sequence of phonemes of finite length. A language is an enormously involved system, and it is quite obvious that any attempt to present directly the set of grammatical phoneme sequences would lead to a grammar so complex that it would be practically useless. For this reason (among others), linguistic description proceeds in terms of a system of "levels of representations." Instead of stating the phonemic structure of sentences directly, the linguist sets up such 'higher level' elements as morphemes, and states separately the morphemic structure of sentences and the phonemic structure of morphemes. It can easily be seen that the joint description of these two levels will be much simpler than a direct description of the phonemic structure of sentences.

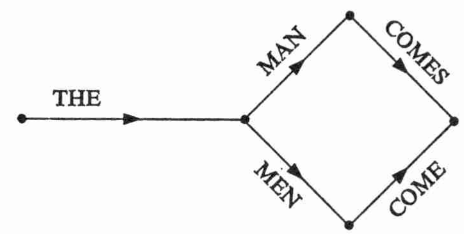
Let us now consider various ways of describing the morphemic structure of sentences. We ask what sort of grammar is necessary to generate all the sequences of morphemes (or words) that constitute grammatical English sentences, and only these.

One requirement that a grammar must certainly meet is that it be finite. Hence the grammar cannot simply be a list of all morpheme (or word) sequences, since there are infinitely many of these. A familiar communication theoretic model for language suggests a way out of this difficulty. Suppose that we have a machine that can be in any one of a finite number of different internal states, and suppose that this machine switches from one state to another by



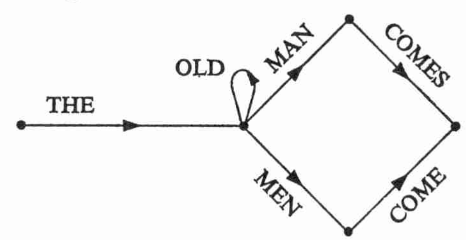
producing a certain symbol (let us say, an English word). One of these states is an *initial state*; another is a *final state*. Suppose that the machine begins in the initial state, runs through a sequence of states (producing a word with each transition), and ends in the final state. Then we call the sequence of words that has been produced a "sentence". Each such machine thus defines a certain language; namely, the set of sentences that can be produced in this way. Any language that can be produced by a machine of this sort we call a *finite state language*; and we can call the machine itself a *finite state grammar*. A finite state grammar can be represented graphically in the form of a "state diagram".[1] For example, the grammar that produces just the two sentences "the man comes" and "the men come" can be represented by the following state diagram:

(7)




We can extend this grammar to produce an infinite number of sentences by adding closed loops. Thus the finite grammar of the subpart of English containing the above sentences in addition to "the old man comes", "the old old man comes", ..., "the old men come", "the old old men come", ..., can be represented by the following state diagram:

(8)




[1] C. E. Shannon and W. Weaver, *The mathematical theory of communication* (Urbana, 1949), pp. 15f.

Given a state diagram, we produce a sentence by tracing a path from the initial point on the left to the final point on the right, always proceeding in the direction of the arrows. Having reached a certain point in the diagram, we can proceed along any path leading from this point, whether or not this path has been traversed before in constructing the sentence in question. Each node in such a diagram thus corresponds to a state of the machine. We can allow transition from one state to another in several ways, and we can have any number of closed loops of any length. The machines that produce languages in this manner are known mathematically as "finite state Markov processes." To complete this elementary communication theoretic model for language, we assign a probability to each transition from state to state. We can then calculate the "uncertainty" associated with each state and we can define the "information content" of the language as the average uncertainty, weighted by the probability of being in the associated states. Since we are studying grammatical, not statistical structure of language here, this generalization does not concern us.

This conception of language is an extremely powerful and general one. If we can adopt it, we can view the speaker as being essentially a machine of the type considered. In producing a sentence, the speaker begins in the initial state, produces the first word of the sentence, thereby switching into a second state which limits the choice of the second word, etc. Each state through which he passes represents the grammatical restrictions that limit the choice of the next word at this point in the utterance.[2]

In view of the generality of this conception of language, and its utility in such related disciplines as communication theory, it is important to inquire into the consequences of adopting this point of view in the syntactic study of some language such as English or a formalized system of mathematics. Any attempt to construct a finite state grammar for English runs into serious difficulties and complications at the very outset, as the reader can easily convince himself. However, it is unnecessary to attempt to show this by

[2] This is essentially the model of language that Hockett develops in *A manual of phonology* (Baltimore, 1955), 02.



example, in view of the following more general remark about English:

(9) English is not a finite state language.

That is, it is *impossible*, not just difficult, to construct a device of the type described above (a diagram such as (7) or (8)) which will produce all and only the grammatical sentences of English. To demonstrate (9) it is necessary to define the syntactic properties of English more precisely. We shall proceed to describe certain syntactic properties of English which indicate that, under any reasonable delimitation of the set of sentences of the language, (9) can be regarded as a theorem concerning English. To go back to the question asked in the second paragraph of § 3, (9) asserts that it is not possible to state the morphemic structure of sentences directly by means of some such device as a state diagram, and that the Markov process conception of language outlined above cannot be accepted, at least for the purposes of grammar.

**3.2** A language is defined by giving its 'alphabet' (i.e., the finite set of symbols out of which its sentences are constructed) and its grammatical sentences. Before investigating English directly, let us consider several languages whose alphabets contain just the letters *a*, *b*, and whose sentences are as defined in (10i–iii):

(10) (i) *ab*, *aabb*, *aaabbb*, ..., and in general, all sentences consisting of *n* occurrences of *a* followed by *n* occurrences of *b* and only these;
(ii) *aa*, *bb*, *abba*, *baab*, *aaaa*, *bbbb*, *aabbaa*, *abbbba*, ..., and in general, all sentences consisting of a string *X* followed by the 'mirror image' of *X* (i.e., *X* in reverse), and only these;
(iii) *aa*, *bb*, *abab*, *baba*, *aaaa*, *bbbb*, *aabaab*, *abbabb*, ..., and in general, all sentences consisting of a string *X* of *a*'s and *b*'s followed by the identical string *X*, and only these.

We can easily show that each of these three languages is not a finite state language. Similarly, languages such as (10) where the *a*'s and *b*'s in question are not consecutive, but are embedded in other strings, will fail to be finite state languages under quite general conditions.[3]

But it is clear that there are subparts of English with the basic form of (10i) and (10ii). Let $S_1$, $S_2$, $S_3$, ... be declarative sentences in English. Then we can have such English sentences as:

(11) (i) If $S_1$, then $S_2$.
(ii) Either $S_3$, or $S_4$.
(iii) The man who said that $S_5$, is arriving today.

In (11i), we cannot have "or" in place of "then"; in (11ii), we cannot have "then" in place of "or"; in (11iii), we cannot have "are" instead of "is". In each of these cases there is a dependency between words on opposite sides of the comma (i.e., "if"–"then", "either"–"or", "man"–"is"). But between the interdependent words, in each case, we can insert a declarative sentence $S_1$, $S_3$, $S_5$, and this declarative sentence may in fact be one of (11i–iii). Thus if in (11i) we take $S_1$ as (11ii) and $S_3$ as (11iii), we will have the sentence:

(12) if, either (11iii), or $S_4$, then $S_2$,

and $S_5$ in (11iii) may again be one of the sentences of (11). It is clear, then, that in English we can find a sequence $a + S_1 + b$, where there is a dependency between $a$ and $b$, and we can select as $S_1$ another sequence containing $c + S_2 + d$, where there is a dependency between $c$ and $d$, then select as $S_2$ another sequence of this form, etc. A set of sentences that is constructed in this way (and we see from (11) that there are several possibilities available for such construction—(11) comes nowhere near exhausting these possibilities) will have all of the mirror image properties of (10ii) which exclude (10ii) from the set of finite state languages. Thus we can find various kinds of non-

[3] See my "Three models for the description of language," *I.R.E. Transactions on Information Theory*, vol. IT-2, Proceedings of the symposium on information theory, Sept., 1956, for a statement of such conditions and a proof of (9). Notice in particular that the set of well-formed formulas of any formalized system of mathematics or logic will fail to constitute a finite state language, because of paired parentheses or equivalent restrictions.



finite state models within English. This is a rough indication of the lines along which a rigorous proof of (9) can be given, on the assumption that such sentences as (11) and (12) belong to English, while sentences that contradict the cited dependencies of (11) (e.g., "either $S_1$, then $S_2$," etc.) do not belong to English. Note that many of the sentences of the form (12), etc., will be quite strange and unusual (they can often be made less strange by replacing "if" by "whenever", "on the assumption that", "if it is the case that", etc., without changing the substance of our remarks). But they are all grammatical sentences, formed by processes of sentence construction so simple and elementary that even the most rudimentary English grammar would contain them. They can be understood, and we can even state quite simply the conditions under which they can be true. It is difficult to conceive of any possible motivation for excluding them from the set of grammatical English sentences. Hence it seems quite clear that no theory of linguistic structure based exclusively on Markov process models and the like, will be able to explain or account for the ability of a speaker of English to produce and understand new utterances, while he rejects other new sequences as not belonging to the language.

**3.3** We might arbitrarily decree that such processes of sentence formation in English as those we are discussing cannot be carried out more than $n$ times, for some fixed $n$. This would of course make English a finite state language, as, for example, would a limitation of English sentences to length of less than a million words. Such arbitrary limitations serve no useful purpose, however. The point is that there are processes of sentence formation that finite state grammars are intrinsically not equipped to handle. If these processes have no finite limit, we can prove the literal inapplicability of this elementary theory. If the processes have a limit, then the construction of a finite state grammar will not be literally out of the question, since it will be possible to list the sentences, and a list is essentially a trivial finite state grammar. But this grammar will be so complex that it will be of little use or interest. In general, the assumption that languages are infinite is made in order to simplify the description of these languages. If a grammar does not have recursive devices (closed loops, as in (8), in the finite state grammar) it will be prohibitively complex. If it does have recursive devices of some sort, it will produce infinitely many sentences.

In short, the approach to the analysis of grammaticalness suggested here in terms of a finite state Markov process that produces sentences from left to right, appears to lead to a dead end just as surely as the proposals rejected in § 2. If a grammar of this type produces all English sentences, it will produce many non-sentences as well. If it produces only English sentences, we can be sure that there will be an infinite number of true sentences, false sentences, reasonable questions, etc., which it simply will not produce.

The conception of grammar which has just been rejected represents in a way the minimal linguistic theory that merits serious consideration. A finite state grammar is the simplest type of grammar which, with a finite amount of apparatus, can generate an infinite number of sentences. We have seen that such a limited linguistic theory is not adequate; we are forced to search for some more powerful type of grammar and some more 'abstract' form of linguistic theory. The notion of "linguistic level of representation" put forth at the outset of this section must be modified and elaborated. At least one linguistic level *cannot* have this simple structure. That is, on some level, it will not be the case that each sentence is represented simply as a finite sequence of elements of some sort, generated from left to right by some simple device. Alternatively, we must give up the hope of finding a *finite* set of levels, ordered from high to low, so constructed that we can generate all utterances by stating the permitted sequences of highest level elements, the constituency of each highest level element in terms of elements of the second level, etc., finally stating the phonemic constituency of elements of the next-to-lowest level.[4] At the outset of § 3, we

[4] A third alternative would be to retain the notion of a linguistic level as a simple linear method of representation, but to generate at least one such level from left to right by a device with more capacity than a finite state Markov process. There are so many difficulties with the notion of linguistic level based on left to right generation, both in terms of complexity of description and lack



proposed that levels be established in this way in order to *simplify* the description of the set of grammatical phoneme sequences. If a language can be described in an elementary, left-to-right manner in terms of a single level (i.e., if it is a finite state language) then this description may indeed be simplified by construction of such higher levels; but to generate non-finite state languages such as English we need fundamentally different methods, and a more general concept of "linguistic level".

of explanatory power (cf. § 8), that it seems pointless to pursue this approach any further. The grammars that we discuss below that do not generate from left to right also correspond to processes less elementary than finite state Markov processes. But they are perhaps less powerful than the kind of device that would be required for direct left-to-right generation of English. Cf. my "Three models for the description of language" for some futher discussion.

# 4

## PHRASE STRUCTURE

**4.1** Customarily, linguistic description on the syntactic level is formulated in terms of constituent analysis (parsing). We now ask what form of grammar is presupposed by description of this sort. We find that the new form of grammar is *essentially* more powerful than the finite state model rejected above, and that the associated concept of "linguistic level" is different in fundamental respects.

As a simple example of the new form for grammars associated with constituent analysis, consider the following:

(13) (i) $Sentence \rightarrow NP + VP$
(ii) $NP \rightarrow T + N$
(iii) $VP \rightarrow Verb + NP$
(iv) $T \rightarrow the$
(v) $N \rightarrow$ *man*, *ball*, etc.
(vi) $Verb \rightarrow$ *hit*, *took*, etc.

Suppose that we interpret each rule $X \rightarrow Y$ of (13) as the instruction "rewrite $X$ as $Y$". We shall call (14) a *derivation* of the sentence "the man hit the ball." where the numbers at the right of each line of the derivation refer to the rule of the "grammar" (13) used in constructing that line from the preceding line.[1]

[1] The numbered rules of English grammar to which reference will constantly be made in the following pages are collected and properly ordered in § 12, *Appendix II*. The notational conventions that we shall use throughout the discussion of English structure are stated in § 11, *Appendix I*.

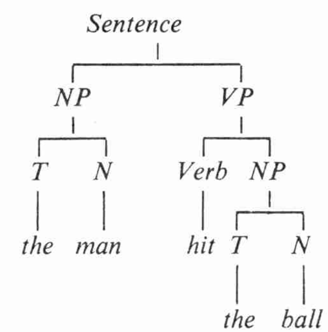
In his "Axiomatic syntax: the construction and evaluation of a syntactic calculus," *Language* 31.409–14 (1955), Harwood describes a system of word class analysis similar in form to the system developed below for phrase structure. The system he describes would be concerned only with the relation between $T + N + Verb + T + N$ and $the + man + hit + the + ball$ in the example discussed



| | | |
|---|---|---|
| (14) | *Sentence* | |
| | $NP + VP$ | (i) |
| | $T + N + VP$ | (ii) |
| | $T + N + Verb + NP$ | (iii) |
| | $the + N + Verb + NP$ | (iv) |
| | $the + man + Verb + NP$ | (v) |
| | $the + man + hit + NP$ | (vi) |
| | $the + man + hit + T + N$ | (ii) |
| | $the + man + hit + the + N$ | (iv) |
| | $the + man + hit + the + ball$ | (v) |

Thus the second line of (14) is formed from the first line by rewriting *Sentence* as $NP + VP$ in accordance with rule (i) of (13); the third line is formed from the second by rewriting $NP$ as $T + N$ in accordance with rule (ii) of (13); etc. We can represent the derivation (14) in an obvious way by means of the following diagram:

(15)




The diagram (15) conveys less information than the derivation (14), since it does not tell us in what order the rules were applied in (14).

in (13)–(15); i.e., the grammar would contain the "initial string" $T + N + Verb + T + N$ and such rules as (13iv–vi). It would thus be a weaker system than the elementary theory discussed in § 3, since it could not generate an infinite language with a finite grammar. While Harwood's formal account (pp. 409–11) deals only with word class analysis, the linguistic application (p. 412) is a case of immediate constituent analysis, with the classes $C_{i..m}$ presumably taken to be classes of word sequences. This extended application is not quite compatible with the formal account, however. For example, none of the proposed measures of goodness of fit can stand without revision under this reinterpretation of the formalism.

Given (14), we can construct (15) uniquely, but not vice versa, since it is possible to construct a derivation that reduces to (15) with a different order of application of the rules. The diagram (15) retains just what is essential in (14) for the determination of the phrase structure (constituent analysis) of the derived sentence "the man hit the ball." A sequence of words of this sentence is a constituent of type $Z$ if we can trace this sequence back to a single point of origin in (15), and this point of origin is labelled $Z$. Thus "hit the ball" can be traced back to $VP$ in (15); hence "hit the ball" is a $VP$ in the derived sentence. But "man hit" cannot be traced back to any single point of origin in (15); hence "man hit" is not a constituent at all.

We say that two derivations are *equivalent* if they reduce to the same diagram of the form (15). Occasionally, a grammar may permit us to construct nonequivalent derivations for a given sentence. Under these circumstances, we say that we have a case of "constructional homonymity",[2] and if our grammar is correct, this sentence of the language should be ambiguous. We return to the important notion of constructional homonymity below.

One generalization of (13) is clearly necessary. We must be able to limit application of a rule to a certain context. Thus $T$ can be rewritten *a* if the following noun is singular, but not if it is plural; similarly, *Verb* can be rewritten "hits" if the preceding noun is *man*, but not if it is *men*. In general, if we wish to limit the rewriting of $X$ as $Y$ to the context $Z - W$, we can state in the grammar the rule

(16) $Z + X + W \rightarrow Z + Y + W.$

For example, in the case of singular and plural verbs, instead of having $Verb \rightarrow hits$ as an additional rule of (13), we should have

(17) $NP_{sing} + Verb \rightarrow NP_{sing} + hits$

indicating that *Verb* is rewritten *hits* only in the context $NP_{sing}-$.

[2] See § 8.1 for some examples of constructional homonymity. See my *The logical structure of linguistic theory* (mimeographed); "Three models for the description of language" (above, p. 22, fn. 3); C. F. Hockett, "Two models of grammatical description," *Linguistics Today*, *Word* 10.210–33 (1954); R. S. Wells, "Immediate constituents," *Language* 23.81–117 (1947) for more detailed discussion.



Correspondingly, (13 ii) will have to be restated to include $NP_{sing}$ and $NP_{pl}$.[3] This is a straightforward generalization of (13). One feature of (13) must be preserved, however, as it is in (17): only a single element can be rewritten in any single rule; i.e., in (16), $X$ must be a single symbol such as $T$, *Verb*, and not a sequence such as $T + N$. If this condition is not met, we will not be able to recover properly the phrase structure of derived sentences from the associated diagrams of the form (15), as we did above.

We can now describe more generally the form of grammar associated with the theory of linguistic structure based upon constituent analysis. Each such grammar is defined by a finite set $\Sigma$ of initial strings and a finite set F of 'instruction formulas' of the form $X \rightarrow Y$ interpreted: "rewrite $X$ as $Y$." Though $X$ need not be a single symbol, only a single symbol of $X$ can be rewritten in forming $Y$. In the grammar (13), the only member of the set $\Sigma$ of initial strings was the single symbol *Sentence*, and F consisted of the rules (i) – (vi); but we might want to extend $\Sigma$ to include, for example, *Declarative Sentence*, *Interrogative Sentence*, as additional symbols. Given the grammar [$\Sigma$, F], we define a *derivation* as a finite sequence of strings, beginning with an initial string of $\Sigma$, and with each string in the sequence being derived from the preceding string by application of one of the instruction formulas of F. Thus (14) is a derivation, and the five-termed sequence of strings consisting of the first five lines of (14) is also a derivation. Certain derivations are *terminated* derivations, in the sense that their final string cannot be rewritten any further by the rules F. Thus (14) is a terminated derivation, but the sequence consisting of the first five

[3] Thus in a more complete grammar, (13 ii) might be replaced by a set of rules that includes the following:

$$NP \rightarrow \begin{Bmatrix} NP_{sing} \\ NP_{pl} \end{Bmatrix}$$
$$NP_{sing} \rightarrow T + N + \emptyset \ (+ \textit{Prepositional Phrase})$$
$$NP_{pl} \rightarrow T + N + S \ (+ \textit{Prepositional Phrase})$$

where $S$ is the morpheme which is singular for verbs and plural for nouns ("comes," "boys"), and $\emptyset$ is the morpheme which is singular for nouns and plural for verbs ("boy," "come"). We shall omit all mention of first and second person throughout this discussion. Identification of the nominal and verbal number affix is actually of questionable validity.

lines of (14) is not. If a string is the last line of a terminated derivation, we say that it is a *terminal* string. Thus *the* + *man* + *hit* + *the* + *ball* is a terminal string from the grammar (13). Some grammars of the form [Σ, F] may have no terminal strings, but we are interested only in grammars that do have terminal strings, i.e., that describe some language. A set of strings is called a *terminal language* if it is the set of terminal strings for some grammar [Σ, F]. Thus each such grammar defines some terminal language (perhaps the 'empty' language containing no sentences), and each terminal language is produced by some grammar of the form [Σ, F]. Given a terminal language and its grammar, we can reconstruct the phrase structure of each sentence of the language (each terminal string of the grammar) by considering the associated diagrams of the form (15), as we saw above. We can also define the grammatical relations in these languages in a formal way in terms of the associated diagrams.

**4.2** In § 3 we considered languages, called "finite state languages", which were generated by finite state Markov processes. Now we are considering terminal languages that are generated by systems of the form [Σ, F]. These two types of languages are related in the following way

*Theorem*: Every finite state language is a terminal language, but there are terminal languages which are not finite state languages.[4]

The import of this theorem is that description in terms of phrase structure is essentially more powerful than description in terms of the elementary theory presented above in § 3. As examples of terminal languages that are not finite state languages we have the languages (10i), (10ii) discussed in § 3. Thus the language (10i), consisting of all and only the strings *ab*, *aabb*, *aaabbb*, ... can be produced by the [Σ, F] grammar (18).

(18) Σ: $Z$
F: $Z \rightarrow ab$
$Z \rightarrow aZb$

[4] See my "Three models for the description of language" (above, p. 22, fn. 3) for proofs of this and related theorems about relative power of grammars.



This grammar has the initial string Z (as (13) has the initial string *Sentence*) and it has two rules. It can easily be seen that each terminated derivation constructed from (18) ends in a string of the language (10i), and that all such strings are produced in this way. Similarly, languages of the form (10ii) can be produced by [Σ, F] grammars (10iii), however, cannot be produced by a grammar of this type, unless the rules embody contextual restrictions.[5]

In § 3 we pointed out that the languages (10i) and (10ii) correspond to subparts of English, and that therefore the finite state Markov process model is not adequate for English. We now see that the phrase structure model does not fail in such cases. We have not proved the adequacy of the phrase structure model, but we have shown that large parts of English which literally cannot be described in terms of the finite-state process model can be described in terms of phrase structure.

Note that in the case of (18), we can say that in the string *aaabbb* of (10i), for example, *ab* is a *Z*, *aabb* is a *Z*, and *aaabbb* itself is a *Z*.[6] Thus this particular string contains three 'phrases,' each of which is a *Z*. This is, of course, a very trivial language. It is important to observe that in describing this language we have introduced a symbol *Z* which is not contained in the sentences of this language. This is the essential fact about phrase structure which gives it its 'abstract' character.

Observe also that in the case of both (13) and (18) (as in every system of phrase structure), each terminal string has many different representations. For example, in the case of (13), the terminal string "the man hit the ball" is represented by the strings *Sentence*, $NP + VP$, $T + N + VP$, and all the other lines of (14), as well as by such strings as $NP + Verb + NP$, $T + N + hit + NP$, which would occur in other derivations equivalent to (14) in the sense there defined. On the level of phrase structure, then, each sentence of the language is represented by a *set* of strings, not by a single string as it

[5] See my "On certain formal properties of grammars", *Information and Control* 2.133–167 (1959).

[6] Where "is a" is the relation defined in § 4.1 in terms of such diagrams as (15).

is on the level of phonemes, morphemes, or words. Thus phrase structure, taken as a linguistic level, has the fundamentally different and nontrivial character which, as we saw in the last paragraph of § 3, is required for some linguistic level. We cannot set up a hierarchy among the various representations of "the man hit the ball"; we cannot subdivide the system of phrase structure into a finite set of levels, ordered from higher to lower, with one representation for each sentence on each of these sublevels. For example, there is no way of ordering the elements *NP* and *VP* relative to one another. Noun phrases are contained within verb phrases, and verb phrases within noun phrases, in English. Phrase structure must be considered as a single level, with a set of representations for each sentence of the language. There is a one-one correspondence between the properly chosen sets of representations, and diagrams of the form (15).

**4.3** Suppose that by a [Σ, F] grammar we can generate all of the grammatical sequences of morphemes of a language. In order to complete the grammar we must state the phonemic structure of these morphemes, so that the grammar will produce the grammatical phoneme sequences of the language. But this statement (which we would call the *morphophonemics* of the language) can also be given by a set of rules of the form "rewrite *X* as *Y*", e.g., for English,

(19) (i) *walk* → /wɔk/
(ii) *take* + *past* → /tuk/
(iii) *hit* + *past* → /hit/
(iv) /...D/ + *past* → /...D/ + /ɪd/ (where D = /t/ or /d/)
(v) /...$C_{unv}$/ + *past* → /...$C_{unv}$/ + /t/ (where $C_{unv}$ is an unvoiced consonant)
(vi) *past* → /d/.
(vii) *take* → /teyk/
etc.

or something similar. Note, incidentally, that order must be defined among these rules — e.g., (ii) must precede (v) or (vii), or we will derive such forms as /teykt/ for the past tense of *take*. In these



morphophonemic rules we need no longer require that only a single symbol be rewritten in each rule.

We can now extend the phrase structure derivations by applying (19), so that we have a unified process for generating phoneme sequence from the initial string *Sentence*. This makes it appear as though the break between the higher level of phrase structure and the lower levels is arbitrary. Actually, the distinction is not arbitrary. For one thing, as we have seen, the formal properties of the rules *X* → *Y* corresponding to phrase structure are different from those of the morphophonemic rules, since in the case of the former we must require that only a single symbol be rewritten. Second, the elements that figure in the rules (19) can be classified into a finite set of levels (e.g., phonemes and morphemes; or, perhaps, phonemes, morphophonemes, and morphemes) each of which is elementary in the sense that a single string of elements of this level is associated with each sentence as its representation on this level (except in cases of homonymity), and each such string represents a single sentence. But the elements that appear in the rules corresponding to phrase structure cannot be classified into higher and lower levels in this way. We shall see below that there is an even more fundamental reason for marking this subdivison into the higher level rules of phrase structure and the lower level rules that convert strings of morphemes into strings of phonemes.

The formal properties of the system of phrase structure make an interesting study, and it is easy to show that further elaboration of the form of grammar is both necessary and possible. Thus it can easily be seen that it would be quite advantageous to order the rules of the set F so that certain of the rules can apply only after others have applied. For example, we should certainly want all rules of the form (17) to apply before any rule which enables us to rewrite *NP* as *NP* + *Preposition* + *NP*, or the like; otherwise the grammar will produce such nonsentences as "the men near the truck begins work at eight." But this elaboration leads to problems that would carry us beyond the scope of this study.

# 5

# LIMITATIONS OF PHRASE STRUCTURE DESCRIPTION

**5.1** We have discussed two models for the structure of language, a communication theoretic model based on a conception of language as a Markov process and corresponding, in a sense, to the minimal linguistic theory, and a phrase structure model based on immediate constituent analysis. We have seen that the first is surely inadequate for the purposes of grammar, and that the second is more powerful than the first, and does not fail in the same way. Of course there are languages (in our general sense) that cannot be described in terms of phrase structure, but I do not know whether or not English is itself literally outside the range of such analysis. However, I think that there are other grounds for rejecting the theory of phrase structure as inadequate for the purpose of linguistic description.

The strongest possible proof of the inadequacy of a linguistic theory is to show that it literally cannot apply to some natural language. A weaker, but perfectly sufficient demonstration of inadequacy would be to show that the theory can apply only clumsily; that is, to show that any grammar that can be constructed in terms of this theory will be extremely complex, *ad hoc*, and 'unrevealing', that certain very simple ways of describing grammatical sentences cannot be accommodated within the associated forms of grammar, and that certain fundamental formal properties of natural language cannot be utilized to simplify grammars. We can gather a good deal of evidence of this sort in favor of the thesis that the form of grammar described above, and the conception of linguistic theory that underlies it, are fundamentally inadequate.

The only way to test the adequacy of our present apparatus is to attempt to apply it directly to the description of English sentences.



As soon as we consider any sentences beyond the simplest type, and in particular, when we attempt to define some order among the rules that produce these sentences, we find that we run into numerous difficulties and complications. To give substance to this claim would require a large expenditure of effort and space, and I can only assert here that this can be shown fairly convincingly.[1] Instead of undertaking this rather arduous and ambitious course here, I shall limit myself to sketching a few simple cases in which considerable improvement is possible over grammars of the form [Σ, F]. In § 8 I shall suggest an independent method of demonstrating the inadequacy of constituent analysis as a means of describing English sentence structure.

**5.2** One of the most productive processes for forming new sentences is the process of conjunction. If we have two sentences $Z + X + W$ and $Z + Y + W$, and if $X$ and $Y$ are actually constituents of these sentences, we can generally form a new sentence $Z - X + and + Y - W$. For example, from the sentences (20a-b) we can form the new sentence (21).

(20) (a) the scene – of the movie – was in Chicago
(b) the scene – of the play – was in Chicago

(21) the scene – of the movie and of the play – was in Chicago.

If $X$ and $Y$ are, however, not constituents, we generally cannot do this.[2] For example we cannot form (23) from (22a-b).

[1] See my *The logical structure of linguistic theory* for detailed analysis of this problem.

[2] (21) and (23) are extreme cases in which there is no question about the possibility of conjunction. There are many less clear cases. For example, it is obvious that "John enjoyed the book and liked the play" (a string of the form $NP - VP + and + VP$) is a perfectly good sentence, but many would question the grammaticalness of, e.g., "John enjoyed and my friend liked the play" (a string of the form $NP + Verb + and + Verb - NP$). The latter sentence, in which conjunction crosses over constituent boundaries, is much less natural than the alternative "John enjoyed the play and my friend liked it", but there is no preferable alternative to the former. Such sentences with conjunction crossing constituent boundaries are also, in general, marked by special phonemic features such as extra long pauses (in our example, between "liked" and "the"), contrastive stress and intonation, failure to reduce vowels and drop final consonants in

(22) (a) the – liner sailed down the – river
(b) the – tugboat chugged up the – river

(23) the – liner sailed down the and tugboat chugged up the – river.

Similarly, if $X$ and $Y$ are both constituents, but are constituents of different kinds (i.e., if in the diagram of the form (15) they each have a single origin, but this origin is labelled differently), then we cannot in general form a new sentence by conjunction. For example, we cannot form (25) from (24a-b).

(24) (a) the scene – of the movie – was in Chicago
(b) the scene – that I wrote – was in Chicago

(25) the scene – of the movie and that I wrote – was in Chicago

In fact, the possibility of conjunction offers one of the best criteria for the initial determination of phrase structure. We can simplify the description of conjunction if we try to set up constituents in such a way that the following rule will hold:

(26) If $S_1$ and $S_2$ are grammatical sentences, and $S_1$ differs from $S_2$ only in that $X$ appears in $S_1$ where $Y$ appears in $S_2$ (i.e., $S_1 = ..X..$ and $S_2 = ..Y..$), and $X$ and $Y$ are constituents of the same type in $S_1$ and $S_2$, respectively, then $S_3$ is a sentence, where $S_3$ is the result of replacing $X$ by $X + and + Y$ in $S_1$ (i.e., $S_3 = ..X + and + Y..$).

rapid speech, etc. Such features normally mark the reading of non-grammatical strings. The most reasonable way to describe this situation would seem to be by a description of the following kind: to form fully grammatical sentences by conjunction, it is necessary to conjoin single constituents; if we conjoin pairs of constituents, and these are major constituents (i.e., 'high up' in the diagram (15)), the resulting sentences are semi-grammatical; the more completely we violate constituent structure by conjunction, the less grammatical is the resulting sentence. This description requires that we generalize the grammatical-ungrammatical dichotomy, developing a notion of degree of grammaticalness. It is immaterial to our discussion, however, whether we decide to exclude such sentences as "John enjoyed and my friend liked the play" as ungrammatical, whether we include them as semi-grammatical, or whether we include them as fully grammatical but with special phonemic features. In any event they form a class of utterances distinct from "John enjoyed the play and liked the book," etc., where constituent structure is preserved perfectly; and our conclusion that the rule for conjunction must make explicit reference to constituent structure therefore stands, since this distinction will have to be pointed out in the grammar.



Even though additional qualification is necessary here, the grammar is enormously simplified if we set up constituents in such a way that (26) holds even approximately. That is, it is easier to state the distribution of "and" by means of qualifications on this rule than to do so directly without such a rule. But we now face the following difficulty: we cannot incorporate the rule (26) or anything like it ir a grammar [Σ, F] of phrase structure, because of certain fundamental limitations on such grammars. The essential property of rule (26) is that in order to apply it to sentences $S_1$ and $S_2$ to form the new sentence $S_3$ we must know not only the actual form of $S_1$ and $S_2$ but also their constituent structure — we must know not only the final shape of these sentences, but also their 'history of derivation.' But each rule $X \to Y$ of the grammar [Σ, F] applies or fails to apply to a given string by virtue of the actual substance of this string. The question of how this string gradually assumed this form is irrelevant. If the string contains $X$ as a substring, the rule $X \to Y$ can apply to it; if not, the rule cannot apply.

We can put this somewhat differently. The grammar [Σ, F] can also be regarded as a very elementary process that generates sentences not from "left to right" but from "top to bottom". Suppose that we have the following grammar of phrase structure:

(27) Σ: *Sentence*
F: $X_1 \to Y_1$
$\vdots$
$X_n \to Y_n$ .

Then we can represent this grammar as a machine with a finite number of internal states, including an initial and a final state. In its initial state it can produce only the element *Sentence*, thereby moving into a new state. It can then produce any string $Y_i$ such that *Sentence* $\to Y_i$ is one of the rules of F in (27), again moving into a new state. Suppose that $Y_i$ is the string ... $X_j$ ... Then the machine can produce the string ... $Y_j$ ... by "applying" the rule $X_j \to Y_j$. The machine proceeds in this way from state to state until it finally produces a terminal string; it is now in the final state. The machine thus produces derivations, in the sense of §4. The important point is that the state of the machine is completely determined by the string it has just produced (i.e., by the last step of the derivation); more specifically, the state is determined by the subset of 'left-hand' elements $X_i$ of F which are contained in this last-produced string. But rule (26) requires a more powerful machine, which can "look back" to earlier strings in the derivation in order to determine how to produce the next step in the derivation.

Rule (26) is also fundamentally new in a different sense. It makes essential reference to two distinct sentences $S_1$ and $S_2$, but in grammars of the [Σ, F] type, there is no way to incorporate such double reference. The fact that rule (26) cannot be incorporated into the grammar of phrase structure indicates that even if this form for grammar is not literally inapplicable to English, it is certainly inadequate in the weaker but sufficient sense considered above. This rule leads to a considerable simplification of the grammar; in fact, it provides one of the best criteria for determining how to set up constituents. We shall see that there are many other rules of the same general type as (26) which play the same dual role.

**5.3** In the grammar (13) we gave only one way of analyzing the element *Verb*, namely, as *hit* (cf. (13vi)). But even with the verbal root fixed (let us say, as *take*), there are many other forms that this element can assume, e.g., *takes*, *has + taken*, *will + take*, *has + been + taken*, *is + being + taken*, etc. The study of these "auxiliary verbs" turns out to be quite crucial in the development of English grammar. We shall see that their behavior is very regular and simply describable when observed from a point of view that is quite different from that developed above. though it appears to be quite complex if we attempt to incorporate these phrases directly into a [Σ, F] grammar.

Consider first the auxiliaries that appear unstressed; for example, "has" in "John has read the book" but not "does" in "John *does* read books."[3] We can state the occurrence of these auxiliaries in declarative sentences by adding to the grammar (13) the following rules:

[3] We return to the stressed auxiliary "do" below, in § 7.1 (45)–(47).



(28) (i) $Verb \rightarrow Aux + V$

(ii) $V \rightarrow$ *hit*, *take*, *walk*, *read*, etc.

(iii) $Aux \rightarrow C(M)\ (have + en)\ (be + ing)\ (be + en)$

(iv) $M \rightarrow$ *will*, *can*, *may*, *shall*, *must*

(29) (i) $C \rightarrow \begin{Bmatrix} S \text{ in the context } NP_{sing}- \\ \emptyset \text{ in the context } NP_{pl}- \\ past \end{Bmatrix}$ [4]

(ii) Let *Af* stand for any of the affixes *past*, *S*, *Ø*, *en*, *ing*. Let *v* stand for any *M* or *V*, or *have* or *be* (i.e., for any non-affix in the phrase *Verb*). Then:

$$Af + v \rightarrow v + Af\ \#,$$

where # is interpreted as word boundary.[5]

(iii) Replace + by # except in the context $v - Af$. Insert # initially and finally.

The interpretation of the notations in (28iii) is as follows: we must choose the element *C*, and we may choose zero or more of the parenthesized elements in the given order. In (29i) we may develop *C* into any of three morphemes, observing the contextual restrictions given. As an example of the application of these rules, we construct a derivation in the style of (14), omitting the initial steps.

(30) *the + man + Verb + the + book* from (13i-v)
*the + man + Aux + V + the + book* (28i)
*the + man + Aux + read + the + book* (28ii)
*the + man + C + have + en + be + ing + read + the + book* (28iii) – we select the elements *C*, *have + en* and *be + ing*.
*the + man + S + have + en + be + ing + read + the + book* (29i)

[4] We assume here that (13ii) has been extended in the manner of fn. 3, above, p. 29, or something similar.

[5] If we were formulating the theory of grammar more carefully, we would interpret # as the concatenation operator on the level of words, while + is the concatention operator on the level of phrase structure. (29) would then be part of the definition of a mapping which carries certain objects on the level of phrase structure (essentially, diagrams of the form (15)) into strings of words. See my *The logical structure of linguistic theory* for a more careful formulation.

$$the + man + have + S \# be + en \# read + ing \# the + book \quad \text{(29ii) – three times.}$$

$$\# the \# man \# have + S \# be + en \# read + ing \# the \# book \# \quad \text{(29iii)}$$

The morphophonemic rules (19), etc., will convert the last line of this derivation into:

(31) the man has been reading the book

in phonemic transcription. Similarly, every other auxiliary verb phrase can be generated. We return later to the question of further restrictions that must be placed on these rules so that only grammatical sequences can be generated. Note, incidentally, that the morphophonemic rules will have to include such rules as: $will + S \rightarrow will$, $will + past \rightarrow would$. These rules can be dropped if we rewrite (28iii) so that either $C$ or $M$, but not both, can be selected. But now the forms *would*, *could*, *might*, *should* must be added to (28iv), and certain 'sequence of tense' statements become more complex. It is immateral to our further discussion which of these alternative analysesis iadopted. Several other minor revisions are possible.

Notice that in order to apply (29i) in (30) we had to use the fact that $the + man$ is a singular noun phrase $NP_{sing}$. That is, we had to refer back to some earlier step in the derivation in order to determine the constituent structure of $the + man$. (The alternative of ordering (29i) and the rule that develops $NP_{sing}$ into $the + man$ in such a way that (29i) must precede the latter is not possible, for a variety of reasons, some of which appear below). Hence, (29i), just like (26), goes beyond the elementary Markovian character of grammars of phrase structure, and cannot be incorporated within the [Σ, F] grammar.

Rule (29ii) violates the requirements of [Σ, F] grammars even more severely. It also requires reference to constituent structure (i.e., past history of derivation) and in addition, we have no way to express the required inversion within the terms of phrase structure. Note that this rule is useful elsewhere in the grammar, at least in the case where *Af* is *ing*. Thus the morphemes *to* and *ing* play a very



similar role within the noun phrase in that they convert verb phrases into noun phrases, giving, e.g.,

(32) $\left\{\begin{matrix}\text{to prove that theorem}\\ \text{proving that theorem}\end{matrix}\right\}$ was difficult.

etc. We can exploit this parallel by adding to the grammar (13) the rule

(33) $NP \rightarrow \left\{\begin{matrix} ing \\ to \end{matrix}\right\} VP$

The rule (29ii) will then convert $ing + prove + that + theorem$ into $proving \# that + theorem$. A more detailed analysis of the $VP$ shows that this parallel extends much further than this, in fact.

The reader can easily determine that to duplicate the effect of (28iii) and (29) without going beyond the bounds of a system [Σ, F] of phrase structure, it would be necessary to give a fairly complex statement. Once again, as in the case of conjunction, we see that significant simplification of the grammar is possible if we are permitted to formulate rules of a more complex type than those that correspond to a system of immediate constituent analysis. By allowing ourselves the freedom of (29ii) we have been able to state the constituency of the auxiliary phrase in (28iii) without regard to the interdependence of its elements, and it is always easier to describe a sequence of independent elements than a sequence of mutually dependent ones. To put the same thing differently, in the auxiliary verb phrase we really have discontinuous elements – e.g., in (30), the elements *have..en* and *be..ing*. But discontinuities cannot be handled within [Σ, F] grammars.[6] In (28iii) we treated these

[6] We might attempt to extend the notions of phrase structure to account for discontinuities. It has been pointed out several times that fairly serious difficulties arise in any systematic attempt to pursue this course. Cf. my "System of syntactic analysis," *Journal of Symbolic Logic* 18.242–56 (1953); C. F. Hockett, "A formal statement of morphemic analysis," *Studies in Linguistics* 10.27–39 (1952); idem, "Two models of grammatical description," *Linguistics Today*, *Word* 10.210–33 (1954). Similarly, one might seek to remedy some of the other deficiencies of [Σ, F] grammars by a more complex account of phrase structure. I think that such an approach is ill-advised, and that it can only lead to the development of *ad hoc* and fruitless elaborations. It appears to be the case that the notions of phrase structure are quite adequate for a small

elements as continuous, and we introduced the discontinuity by the very simple additional rule (29 ii). We shall see below, in § 7, that this analysis of the element *Verb* serves as the basis for a far-reaching and extremely simple analysis of several important features of English syntax.

**5.4** As a third example of the inadequacy of the conceptions of phrase structure, consider the case of the active-passive relation. Passive sentences are formed by selecting the element *be + en* in rule (28 iii). But there are heavy restrictions on this element that make it unique among the elements of the auxiliary phrase. For one thing, *be + en* can be selected only if the following *V* is transitive (e.g., *was + eaten* is permitted, but not *was + occurred*); but with a few exceptions the other elements of the auxiliary phrase can occur freely with verbs. Furthermore, *be + en* cannot be selected if the verb *V* is followed by a noun phrase, as in (30) (e.g., we cannot in general have *NP + is + V + en + NP*, even when *V* is transitive – we cannot have "lunch is eaten John"). Furthermore, if *V* is transitive and is followed by the prepositional phrase *by + NP*, then we *must* select *be + en* (we can have "lunch is eaten by John" but not "John is eating by lunch," etc.). Finally, note that in elaborating (13) into a full-fledged grammar we will have to place many restrictions on the choice of *V* in terms of subject and object in order to permit such sentences as: "John admires sincerity," "sincerity frightens John," "John plays golf," "John drinks wine," while excluding the 'inverse' non-sentences[7] "sincerity admires John," "John frightens sincerity," "golf plays John," "wine drinks John". But this whole network of restrictions fails completely when we choose *be + en* as part of the auxiliary verb. In fact, in this case the same selectional dependencies hold, but in the opposite order. That is, for every sentence $NP_1 - V - NP_2$ we can have a corresponding sentence $NP_2 - is + Ven - by + NP_1$. If we try to include passives directly in the grammar (13), we shall have to restate all of these restrictions in the opposite order for the case in which *be + en* is chosen as part of the auxiliary verb. This inelegant duplication, as well as the special restrictions involving the element *be + en*, can be avoided only if we deliberately exclude passives from the grammar of phrase structure, and reintroduce them by a rule such as:

(34) If $S_1$ is a grammatical sentence of the form

$$NP_1 - Aux - V - NP_2,$$

then the corresponding string of the form

$$NP_2 - Aux + be + en - V - by + NP_1$$

is also a grammatical sentence.

For example, if *John – C – admire – sincerity* is a sentence, then *sincerity – C + be + en – admire – by + John* (which by (29) and (19) becomes "sincerity is admired by John") is also a sentence.

We can now drop the element *be + en*, and all of the special restrictions associated with it, from (28 iii). The fact that *be + en* requires a transitive verb, that it cannot occur before *V + NP*, that it must occur before *V + by + NP* (where *V* is transitive), that it inverts the order of the surrounding noun phrases, is in each case an automatic consequence of rule (34). This rule thus leads to a considerable simplification of the grammar. But (34) is well beyond the limits of [Σ, F] grammars. Like (29 ii), it requires reference to the constituent structure of the string to which it applies and it carries out an inversion on this string in a structurally determined manner.

---

part of the language and that the rest of the language can be derived by repeated application of a rather simple set of transformations to the strings given by the phrase structure grammar. If we were to attempt to extend phrase structure grammar to cover the entire language directly, we would lose the simplicity of the limited phrase structure grammar and of the transformational development. This approach would miss the main point of level construction (cf. first paragraph of § 3.1), namely, to rebuild the vast complexity of the actual language more elegantly and systematically by extracting the contribution to this complexity of several linguistic levels, each of which is simple in itself.

[7] Here too we might make use of a notion of levels of grammaticalness as suggested in footnote 2, p. 35. Thus "sincerity admires John," though clearly less grammatical than "John admires sincerity," is certainly more grammatical than "of admires John," I believe that a workable notion of degree of grammaticalness can be developed in purely formal terms (cf. my *The logical structure of linguistic theory*), but this goes beyond the bounds of the present discussion. See § 7.5 for an even stronger demonstration that inversion is necessary in the passive.

**5.5** We have discussed three rules ((26), (29), (34)) which materially simplify the description of English but which cannot be incorporated into a [Σ, F] grammar. There are a great many other rules of this type, a few of which we shall discuss below. By further study of the limitations of phrase structure grammars with respect to English we can show quite conclusively that these grammars will be so hopelessly complex that they will be without interest unless we incorporate such rules.

If we examine carefully the implications of these supplementary rules, however, we see that they lead to an entirely new conception of linguistic structure. Let us call each such rule a "grammatical transformation." A grammatical transformation T operates on a given string (or, as in the case of (26), on a set of strings) with a given constituent structure and converts it into a new string with a new derived constituent structure. To show exactly *how* this operation is performed requires a rather elaborate study which would go far beyond the scope of these remarks, but we can in fact develop a certain fairly complex but reasonably natural algebra of transformations having the properties that we apparently require for grammatical description.[8]

From these few examples we can already detect some of the essential properties of a transformational grammar. For one thing, it is clear that we must define an order of application on these transformations. The passive transformation (34), for example, must apply *before* (29). It must precede (29i), in particular, so that the verbal element in the resulting sentence will have the same number as the new grammatical subject of the passive sentence. And it must precede (29ii) so that the latter rule will apply properly to the new inserted element *be + en*. (In discussing the question of whether or not (29i) can be fitted into a [Σ, F] grammar, we mentioned that this rule could not be required to apply before the rule

[8] See my "Three models for the description of language" (above, p. 22, fn. 3) for a brief account of transformations, and *The logical structure of linguistic theory* and *Transformational Analysis* for a detailed development of transformational algebra and transformational grammars. See Z. S. Harris, "Cooccurrence and Transformations in linguistic structure," *Language* 33.283-340 (1957), for a somewhat different approach to transformational analysis.



analyzing $NP_{sing}$ into *the + man*, etc. One reason for this is now obvious — (29i) must apply after (34), but (34) must apply after the analysis of $NP_{sing}$, or we will not have the proper selectional relations between the subject and verb and the verb and 'agent' in the passive.)

Secondly, note that certain transformations are *obligatory*, whereas others are only *optional*. For example, (29) must be applied to every derivation, or the result will simply not be a sentence.[9] But (34), the passive transformation, may or may not be applied in any particular case. Either way the result is a sentence. Hence (29) is an obligatory transformation and (34) is an optional transformation.

This distinction between obligatory and optional transformations leads us to set up a fundamental distinction among the sentences of the language. Suppose that we have a grammar G with a [Σ, F] part and a transformational part, and suppose that the transformational part has certain obligatory transformations and certain optional ones. Then we define the *kernel* of the language (in terms of the grammar G) as the set of sentences that are produced when we apply obligatory transformations to the terminal strings of the [Σ, F] grammar. The transformational part of the grammar will be set up in such a way that transformations can apply to kernel sentences (more correctly, to the forms that underlie kernel sentences—i.e., to terminal strings of the [Σ, F] part of the grammar) or to prior transforms. Thus every sentence of the language will either belong to the kernel or will be derived from the strings underlying one or more kernel sentences by a sequence of one or more transformations.

From these considerations we are led to a picture of grammars as possessing a natural tripartite arrangement. Corresponding to the level of phrase structure, a grammar has a sequence of rules of the form $X \rightarrow Y$, and corresponding to lower levels it has a sequence of

[9] But of the three parts of (29i), only the third is obligatory. That is, *past* may occur after $NP_{sing}$ or $NP_{pl}$. Whenever we have an element such as *C* in (29i) which must be developed, but perhaps in several alternative ways, we can order the alternatives and make each one but the last optional, and the last, obligatory.

morphophonemic rules of the same basic form. Linking these two sequences, it has a sequence of transformational rules. Thus the grammar will look something like this:

(35)
$$\begin{array}{ll}
\Sigma: & \textit{Sentence}: \\
\mathrm{F}: & \left.\begin{array}{l} X_1 \rightarrow Y_1 \\ \quad\vdots \\ X_n \rightarrow Y_n \end{array}\right\} \text{Phrase structure} \\
 & \left.\begin{array}{l} \mathrm{T}_1 \\ \quad\vdots \\ \mathrm{T}_j \end{array}\right\} \text{Transformational structure} \\
 & \left.\begin{array}{l} Z_1 \rightarrow W_1 \\ \quad\vdots \\ Z_m \rightarrow W_m \end{array}\right\} \text{Morphophonemics}
\end{array}$$

To produce a sentence from such a grammar we construct an extended derivation beginning with *Sentence*. Running through the rules of F we construct a terminal string that will be a sequence of morphemes, though not necessarily in the correct order. We then run through the sequence of transformations $\mathrm{T}_1, \dots \mathrm{T}_j$, applying each obligatory one and perhaps certain optional ones. These transformations may rearrange strings or may add or delete morphemes. As a result they yield a string of words. We then run through the morphophonemic rules, thereby converting this string of words into a string of phonemes. The phrase structure segment of the grammar will include such rules as those of (13), (17) and (28). The transformational part will include such rules as (26), (29) and (34), formulated properly in the terms that must be developed in a full-scale theory of transformations. The morphophonemic part will include such rules as (19). This sketch of the process of generation of sentences must (and easily can) be generalized to allow for proper functioning of such rules as (26) which operate on a set of sentences, and to allow transformations to reapply to transforms so that more and more complex sentences can be produced.

When we apply only obligatory transformations in the generation of a given sentence, we call the resulting sentence a kernel sentence. Further investigation would show that in the phrase structure and



morphophonemic parts of the grammar we can also extract a skeleton of obligatory rules that *must* be applied whenever we reach them in the process of generating a sentence. In the last few paragraphs of § 4 we pointed out that the phrase structure rules lead to a conception of linguistic structure and "level of representation" that is fundamentally different from that provided by the morphophonemic rules. On each of the lower levels corresponding to the lower third of the grammar an utterance is, in general, represented by a single sequence of elements. But phrase structure cannot be broken down into sublevels: on the level of phrase structure an utterance is represented by a set of strings that cannot be ordered into higher or lower levels. This set of representing strings is equivalent to a diagram of the form (15). On the transformational level, an utterance is represented even more abstractly in terms of a sequence of transformations by which it is derived, ultimately from kernel sentences (more correctly, from the strings which underlie kernel sentences). There is a very natural general definition of "linguistic level" that includes all of these cases,[10] and as we shall see later, there is good reason to consider each of these structures to be a linguistic level.

When transformational analysis is properly formulated we find that it is essentially more powerful than description in terms of phrase structure, just as the latter is essentially more powerfull than description in terms of finite state Markov processes that generate sentences from left to right. In particular, such languages as (10iii) which lie beyond the bounds of phrase structure description with context-free rules can be derived transformationally.[11] It is important to observe that the grammar is materially simplified when we add a transformational level, since it is now necessary to provide phrase structure directly only for kernel sentences — the terminal strings of the [Σ, F] grammar are just those which underlie kernel

[10] Cf. *The logical structure of linguistic theory* and *Transformational Analysis*.

[11] Let G be a [Σ, F] grammar with the initial string *Sentence* and with the set of all finite strings of *a*'s and *b*'s as its terminal output. There is such a grammar. Let G′ be the grammar which contains G as its phrase structure part, supplemented by the transformation T that operates on any string $K$ which is a *Sentence*, converting it into $K + K$. Then the output of G′ is (10iii). Cf. p. 31.

sentences. We choose the kernel sentences in such a way that the terminal strings underlying the kernel are easily derived by means of a [Σ, F] description, while all other sentences can be derived from these terminal strings by simply statable transformations. We have seen, and shall see again below, several examples of simplifications resulting from transformational analysis. Full-scale syntactic investigation of English provides a great many more cases.

One further point about grammars of the form (35) deserves mention, since it has apparently led to some misunderstanding. We have described these grammars as devices for generating sentences. This formulation has occasionally led to the idea that there is a certain asymmetry in grammatical theory in the sense that grammar is taking the point of view of the speaker rather than the hearer; that it is concerned with the process of producing utterances rather than the 'inverse' process of analyzing and reconstructing the structure of given utterances. Actually, grammars of the form that we have been discussing are quite neutral as between speaker and hearer, between synthesis and analysis of utterances. A grammar does not tell us how to synthesize a specific utterance; it does not tell us how to analyze a particular given utterance. In fact, these two tasks which the speaker and hearer must perform are essentially the same, and are both outside the scope of grammars of the form (35). Each such grammar is simply a description of a certain set of utterances, namely, those which it generates. From this grammar we can reconstruct the formal relations that hold among these utterances in terms of the notions of phrase structure, transformational structure, etc. Perhaps the issue can be clarified by an analogy to a part of chemical theory concerned with the structurally possible compounds. This theory might be said to generate all physically possible compounds just as a grammar generates all grammatically 'possible' utterances. It would serve as a theoretical basis for techniques of qualitative analysis and synthesis of specific compounds, just as one might rely on a grammar in the investigation of such special problems as analysis and synthesis of particular utterances.

# 6

# ON THE GOALS OF LINGUISTIC THEORY

**6.1** In §§ 3, 4 two models of linguistic structure were developed: a simple communication theoretic model and a formalized version of immediate constituent analysis. Each was found to be inadequate, and in § 5 I suggested a more powerful model combining phrase structure and grammatical transformations that might remedy these inadequacies. Before going on to explore this possiblity, I would like to clarify certain points of view that underlie the whole approach of his study.

Our fundamental concern throughout this discussion of linguistic structure is the problem of justification of grammars. A grammar of the language L is essentially a theory of L. Any scientific theory is based on a finite number of observations, and it seeks to relate the observed phenomena and to predict new phenomena by constructing general laws in terms of hypothetical constructs such as (in physics, for example) "mass" and "electron." Similarly, a grammar of English is based on a finite corpus of utterances (observations), and it will contain certain grammatical rules (laws) stated in terms of the particular phonemes, phrases, etc., of English (hypothetical constructs). These rules express structural relations among the sentences of the corpus and the indefinite number of sentences generated by the grammar beyond the corpus (predictions). Our problem is to develop and clarify the criteria for selecting the correct grammar for each language, that is, the correct theory of this language.

Two types of criteria were mentioned in § 2.1. Clearly, every grammar will have to meet certain *external conditions of adequacy*; e.g., the sentences generated will have to be acceptable to the native