

Linguistic Society of America

Review: [untitled]

Author(s): Robert B. Lees

Reviewed work(s):

Syntactic Structures by Noam Chomsky

Source: *Language*, Vol. 33, No. 3, Part 1 (Jul. - Sep., 1957), pp. 375-408

Published by: Linguistic Society of America

Stable URL: <http://www.jstor.org/stable/411160>

Accessed: 14/09/2009 11:31

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=lsa>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



Linguistic Society of America is collaborating with JSTOR to digitize, preserve and extend access to *Language*.

<http://www.jstor.org>

REVIEWS

Syntactic structures. By NOAM CHOMSKY. Pp. 116. 's-Gravenhage: Mouton & Co., N.V., 1957.

Reviewed by ROBERT B. LEES, *Massachusetts Institute of Technology*

During the past twenty-five years of development in the field of linguistics it has become customary, especially among anthropologists, to regard linguistics as a very advanced, systematic, precise, powerful—in short, 'scientific' discipline. Social scientists, seeking greater precision in the formulation of their special theories, look with envy upon the very exact grammatical statements and the impressive laboratory-phonetic descriptions which are to be found in the pages of linguistic journals, and the linguist himself has come to believe in some instances that he alone among social scientists points the way to a new scientific revolution in our understanding of human behavior.

Before considering in detail the contents of Chomsky's monograph it is useful to re-examine our notions of science and linguistics, for this little book on syntactic structure has much to say about the status of linguistics as a science.¹ Admitting that that typically western institution which we call science is marked by a number of different kinds of activity, such as the use of precise measurement, complicated technical devices, laboratory experiments, or statistical analysis of masses of data, such activities as these are not nearly so characteristic of scientific method as is theory construction and validation. This may be seen most clearly when we compare in the overall history of one of the physical sciences the various stages of achievement through which the science has passed.

Consider as an example the development of chemistry as we know it today. While some chemical knowledge is very old indeed, all that the ancient world knew of chemistry can be considered a kind of accumulated tradition of folk lore, the rules of thumb passed on from one generation of embalmers, tanners, and

¹ The monograph reviewed here is a version of some parts abstracted from a much larger work now in preparation, together with summaries of other materials previously published and unpublished. The reviewer was privileged to read a first version of the larger work—*The logical structure of linguistic theory*, pp. 752 + li (Cambridge, Mass., 1956)—and now finds it difficult to refrain from referring to topics and results which appear there but not in the book under review. This discussion may therefore serve in part as a preview of the author's detailed statement of his theory of language.

The reader will profit also from three other publications of Chomsky's: Systems of syntactic analysis, *Journal of symbolic logic* 18.242-56 (1953); Semantic considerations in grammar, *Report of the sixth annual Round Table Meeting on Linguistics and Language Teaching* (= *Monograph series on languages and linguistics*, No. 8) 141-50 (1955); Three models for the description of language, *IRE transactions on information theory*, Vol. IT-2, No. 3 (1956). See also his doctoral dissertation, *Transformational analysis* (MS; University of Pennsylvania, 1955).

The reviewer is a member of the Research Laboratory of Electronics, Massachusetts Institute of Technology, where his work is supported in part by the U. S. Army (Signal Corps), the U. S. Air Force (Office of Scientific Research, Air Research and Development Command), and the U. S. Navy (Office of Naval Research), and in part by the National Science Foundation.

cosmeticians to the next, a kind of practical engineering for the householder and artisan. With medieval alchemical study, though it was under the influence of gold-seeking and magic, began the first systematic recording of chemical properties, and, along with the careful observations, the first clumsy theories of chemical behavior, culminating in the rather fanciful but seriously intended phlogiston theory of combustion. With the proposal of this erroneous theory and its subsequent demise in favor of Lavoisier's more satisfactory oxygen theory of burning, alchemical lore graduated from magical engineering to scientific discipline. It is not so much the correctness of Lavoisier's explanation which ushered in the scientific era for chemistry, as the change in attitude which accompanied the whole controversy over combustion theories. From that controversy on, chemistry has been concerned less with observation and classification of curious reactions and more and more with the proposal and validation of theories to account for chemical behavior.

Once it has developed beyond the prescientific stage of collection and classification of interesting facts, a scientific discipline is characterized essentially by the introduction of abstract constructs in theories and the validation of those theories by testing their predictive power. Scientific theories are marked not only by freedom from internal contradiction, but also by maximal cohesion with other branches of knowledge, maximal validity in coverage of known data, and maximal elegance of statement. Freedom from internal contradiction is required for making predictions, for from contradictory statements any assertion may be deduced. Cohesion with other theories is required so that contradictions will not appear in the areas of overlap of data. Maximal coverage is desirable because it is the very purpose of scientific theories to explain by means of generalizations our understanding of particular events and our intuitive perception of their interrelations. Elegance or simplicity of statement compensates for the inevitable limitations on validation which are entailed by finite size of the corpus of data available to us at any time; it may be thought of as a measure of the degree of generality achieved. When scientific prediction diverges from our prescientific, intuitive perception of empirical relations, either the theory is at fault or our perception in error; but when they coincide, the proposed theory is thereby greatly strengthened.

The term 'intuition' may also be used in a different sense in connection with linguistic theories. The empirical data which a linguistic theory must explain consists not only of the noises which talking people produce, but also of various kinds of judgments they can make and feelings they may have about linguistic data. For example, a speaker knows which sounds or words rime, and he can say when a sentence is grammatically permitted or excluded, whether or not he has ever before heard the particular rime or sentence in question. These judgments are sometimes referred to as linguistic intuitions. They are part of the linguistic data to be accounted for, and they must be distinguished from the intuitive or prescientific perceptions which the linguist, qua scientist, has about the data, and which he renders explicit or formalizes, and thus eliminates, by means of his linguistic theory.

While some areas of systematic knowledge are still characterized largely by

minute observations and cataloging, but as yet by little useful generalization, for example in the areas of psychology, economics, medicine, and sociology, some fields of knowledge have reached such an advanced stage of development that their basic problems can be stated very succinctly, and their structure is so well understood that we can now begin to construct axiomatic theories to render explicitly and rigorously just what their content is and means. Thus, we might formulate the fundamental problem of physical chemistry as follows: given the detailed nuclear and electronic structure of an atom, what will be the predicted chemical properties of that substance?

Linguistics and science. How does linguistic science measure up against this conception of scientific theory? Is it possible to state exactly the fundamental problem of linguistics? Confining our attention to that branch of language studies dealing with grammar—grammar in the broad sense: the study of linguistic structures as a whole—at what stage of development are our ideas on grammar? Do we deal with grammatical theories, and if so, are those theories characterized by consistency, cohesion with other branches of behavioral science, elegance of statement, i.e. powerful generality, and compatibility with linguistic intuition? These are serious questions which we must all face honestly.

It would seem that our conception of what the grammar of a language is like is all too often of the purely taxonomic, data-cataloging sort. When we compare a modern descriptive grammar with an old-fashioned prescriptive grammar of a century ago, we are accustomed to dismiss the latter as unscientific, especially to the extent that it slavishly reproduces Latin and Greek grammatical categories in an effort to order the data of a non-classical language. But what more is our descriptive grammar than another reordering of the data—now, to be sure, according to a less traditional scheme of categories, but nonetheless according to an arbitrary set of descriptive labels which has become fossilized within linguistic tradition? Thus, without giving any internal linguistic justification, no reasons derived from a theory of language structure and behavior, the empirical data are organized in our descriptive grammar into chapters on Phonemes, Morphophonemics, Word-formation, The Noun, The Verb, Particles, and possibly Syntax, the whole intended from the very beginning to be just a classification of utterance fractions so that they may be successively mentioned from the first to the last page of the grammar in some manner other than randomly.²

This long digression into the position of grammar among the sciences seemed appropriate here because, in the reviewer's opinion, Chomsky's book on syntactic structures is one of the first serious attempts on the part of a linguist to construct within the tradition of scientific theory-construction a comprehensive theory of language which may be understood in the same sense that a chemical, biological theory is ordinarily understood by experts in those fields. It is not a mere

² It is interesting to note that Hayward Keniston, who describes himself as a humanist, criticizes linguistics as not speculative enough and too bound to a description of physical entities; see *Monograph series on languages and linguistics* 7.146-52 (1954). While criticisms of this sort on the part of humanists may often be interpreted to mean something like 'Linguistics is too scientific, not vague enough', it would be far healthier for us to take it quite literally to mean, 'Linguistics is not scientific enough, too atomistic and taxonomic.'

reorganization of the data into a new kind of library catalog, nor another speculative philosophy about the nature of Man and Language, but rather a rigorous explication of our intuitions about our language in terms of an overt axiom system, the theorems derivable from it, explicit results which may be compared with new data and other intuitions, all based plainly on an overt theory of the internal structure of languages; and it may well provide an opportunity for the application of explicit measures of simplicity to decide preference of one form over another form of grammar.

Discovery, decision, or evaluation? Before going on to examine particularities it is necessary to clarify one essential issue which may well prove to be an obstacle to understanding for many a linguist reader. Although Chomsky discusses this point at some length (§6.1), it is so important that it can easily bear repetition here.

Nearly all American scholarly work in the field of descriptive and structural linguistics is strongly and unmistakably oriented toward the problem of providing linguistics with a mechanical procedure for discovering the correct grammar of any given language, or better, of some given corpus.³

Let us consider first this latter, subsidiary notion which some have entertained, namely, that the linguist has discharged his task once he has specified exactly the constituency of some previously selected, presumably representative text. This criticism is intended to apply equally well to the study of extinct languages, for while the corpus of materials available at any one time for a contemporary language may be vastly larger than the few extant inscriptions for the extinct one, it is nevertheless still finite. Now, for any finite set of objects, say phonemes or words, there are innumerable ways to specify exactly the content and arrangement of those elements, and if this were all that a grammar had to do, there would indeed be a completely mechanical and trivial way to discover grammar. But when we require of the grammar, as does any scientist, that the statements in it be maximally general, it is because we suppose that in this way these statements may be used to predict correctly a maximum number of new elements and sequences which do NOT occur in the finite corpus of data with which we start. Thus we may reasonably expect of a grammar that it account not merely for the sentences of the text in hand, but also for many other sentences which have never been uttered before, but which could be uttered naturally by a speaker of the language in question. There is no obvious mechanical way to generate such a maximally general set of statements.

Time and time again a grammatical proposal is criticised in our literature with the questions: how do you know where to draw that boundary in the stream of

³ Pursuit of this goal is seen in perhaps its best and most resolute form in the works of Zellig S. Harris: *Discontinuous morphemes*, *Lg.* 21.121-7 (1945); *Morpheme to utterance*, *Lg.* 22.161-83 (1946); *Methods in structural linguistics* (Chicago, 1951); *Discourse analysis*, *Lg.* 28.1-30 (1952); *Discourse analysis: A sample text*, *Lg.* 28.474-94 (1952); *Distributional structure*, *Word* 10.146-62 (1954); *Phoneme to morpheme*, *Lg.* 31.190-222 (1955); *Co-occurrence and transformation in linguistic structure*, in the same number of *Language* as this review. Following Harris, Rulon Wells attempted to provide a firm basis for discovery of grammatical structure: *Immediate constituents*, *Lg.* 23.81-117 (1947); but he denied there (93) that it would be an entirely mechanical procedure.

speech? what airtight method have you given us by means of which we may transform our transcriptions or tape-recordings into strings of phonemes or grammatical formulas? how did you find these syntactic categories before having analyzed the morphology, or these stresses before those junctures? It is asked, in other words, that the linguist specify just how the elements which make up the grammar of a language have been elicited from Nature, how they have been deduced from the physical data; this is very much as if it had been demanded of Kekulé that for the success of his theory of carbon compounds he specify exactly how he fixed upon the notion of the six-membered benzene ring. We must take full note here of just what is being required of linguistic theory when we desire that it provide us with a mechanical means for constructing the grammar of a given language, that it provide us, in other words, with a 'discovery procedure'.

It is necessary in this connection to distinguish clearly between this question of grammatical criteria, motivation, and aims on the one hand, and on the other, the independent and trivial question of the physical arrangement or presentation of the grammatical description, once it has been constructed.⁴

Not even the most advanced of the physical sciences, not to mention the whole remaining less exact body of scientific knowledge, is so powerful as to provide a discovery procedure for its area of interest. There is no known mechanical procedure in all of advanced theoretical physics which will permit an expert physicist to find the laws of nature which connect the readings on the meters of his laboratory one with another or each with the phenomena outside of the laboratory. It is a common misconception on the part of many a scientist, strange to say, that correct scientific theories are discovered by making many observations of nature, that somehow the right answers just leap up out of the laboratory notebook if only we have measured enough things accurately.

This is not to say that the scientist can neglect his proficiency in the use of well established and effective laboratory techniques or that the linguist can afford to omit learning how to deal with the informant and collect useful data. Every scientist uses as a source for inspiration in the construction of new models for nature all that he has learned about older successful models and all that he can learn about model construction from older experienced scientists.

American linguistic canons are particularly characterized by this confusion of field and laboratory techniques for data collection and classification on the one hand and model construction or grammar writing on the other. The confusion is further compounded by the often-heard suggestion that statistical methods, that is, elaborate counting techniques, will not only reveal the correct analysis but even explicate linguistic behavior.⁵ Statistical methods ARE in a sense mechanical, nearly everyone knows how to count, they are amenable to machine manipulation in electronic computers, and every statistical formula, even if misapplied, will yield a concrete number, a statistic; but, though much material may be thereby

⁴ Wells, Immediate constituents §46.

⁵ Charles F. Hockett, in a review of Shannon and Weaver's *The mathematical theory of communication*, *Lg.* 29.69-93, esp. 87-8 (1953), proposed that morphemic segmentation might be mechanized by a statistical technique. Similar suggestions are attempted by Harris, Phoneme to morpheme, and by Seymour B. Chatman, Immediate constituents and expansion analysis, *Word* 11.377-85 (1955).

easily summarized, it is not thereby explained. Recent suggestions that phonemic and morphemic segmentation be mechanized by a statistical technique are best regarded as devices for generating hypotheses about linguistic boundaries which must then be validated grammatically. There does not seem to be any demonstrable connection between the grammatical significance of a form and its relative frequency of occurrence, as some statistical theories assume; in fact, it is not even clear that the latter frequency exists as a definable quantity.

The linguistic units postulated by a grammar are constructed much like the concepts of proton, covalent bond, or gene: they are postulated because of the great predictive power which they lend to the theory, but they are not brought to light in the data by a process of induction. And the theories by means of which we order our experiences, on the street or in the laboratory, are generated only by those flashes of insight, those perceptions of pattern, which mark off the brilliant scientist from the dull cataloger of data.⁶

If then we cannot require of linguistic meta-theory that it provide a discovery procedure for linguistics, a procedure for generating grammars mechanically, perhaps one might ask for the next best thing, a mechanical procedure for recognizing a correct grammar (from among all the possible alternatives: i.e. we do the generating, the machine rejects all the wrong ones). Again this would be a rather extravagant request, for even such a well-grounded discipline as classical mathematics cannot boast a 'decision procedure' which would determine mechanically whether any given statement is a theorem or not (except in the most trivial logical calculi). Therefore it does not seem at all over-modest to seek important results on the basis of still weaker demands.

Thus, we might relax our requirements for a linguistic meta-theory to the weaker position of what Chomsky has termed an "evaluation procedure", a mechanical way to evaluate two proposed grammars on the basis of explicit criteria of excellence and reject one as inferior to the other. It would then no longer be the responsibility of the grammarian to state rigorously how he managed to find the particular grammar proposed. Any manipulatory or heuristic principles or devices which he may have found useful or stimulating play no role within the theory itself, once it is constructed.

It may seem strange to some linguists that a grammar can be considered to be a theory of a particular language, and not just a reordering or abbreviation of a text (§6.1). But when we consider the generality which must be required of a grammar, in order that it permit the prediction of an unbounded number of new sentences, just as any speaker himself can generate them, we see that it is analogous not to a herbarium, anatomical map, or library catalog, but rather to a scientific theory embodying proposed laws of nature.

Now, although no natural science can claim even a true evaluation procedure, Chomsky guesses that linguistics might very well permit this degree of theoretical power, and that we shall be in a much better position to determine this if we would only give up our ambitious attempts to provide a discovery or decision procedure for our grammars. To reject the worse alternative of a pair of proposed theories,

⁶ Cf. Carl G. Hempel, *Fundamentals of concept formation* 36-7 (International encyclopedia of unified science 2.7; Chicago, 1952).

the best that the natural scientist can do is to propose a so-called 'crucial experiment'; but linguistics may be able to go one step further and formulate rigorous criteria of excellence of grammars. Chomsky, in the course of axiomatizing grammatical structure, has proposed some ways in which this might be done in the future. For example, with a satisfactory and comprehensive linguistic meta-theory, the notation for structural statements might be standardized, and then the number of symbols used in a grammar would be a measure of its generality.

The belief that linguistics is a discovery procedure for grammars has led quite naturally to the uncritical acceptance of another troublesome assumption, that of phonemic bi-uniqueness. By this is meant that for a transcription to be phonemic, it is necessary that it be unique in two directions: not only must each string of phonemic symbols be pronounceable in one and only one way, but also every utterance must be transcribable in one and only one way in terms of the phonemes. This latter condition of unique transcribability not only is superfluous for linguistics, it does not even render properly the desired condition of natural, automatic transcribability (for every sound type in every environment there should be some one natural transcription), since bi-uniqueness may be achieved in any number of trivial ways. One need only set up a scheme to reject all but some one possible transcription for any utterance, say the first one in any ordered list of all possible ones.⁷

Levels. There is an important and immediate result from our agreement to surrender the requirement for a discovery procedure, namely that much of the motivation for a strict isolation of grammatical levels is thereby lost,⁸ i.e. there is no longer any strong reason to insist that phonemic description (say) be entirely independent of and antecedent to syntactic description. Now there can be no charge of circularity in the use of syntactic criteria to isolate phonemic entities which must in turn be used later on to determine those very syntactic units used.⁹ If one does not demand a mechanical procedure for deriving a gram-

⁷ See below for a case in which superfluous rules can be eliminated by rejecting this principle of bi-uniqueness. Cf. Bernard Bloch, Phonemic overlapping, *Amer. speech* 16.278-84 (1941); Hockett, review of Martinet's *Phonology as functional phonetics*, *Lg.* 27.340 (1951). Chomsky, Halle, and Lukoff have chosen to reject the principle of bi-uniqueness entirely in their article On accent and juncture in English, *For Roman Jakobson* 65-80 (The Hague, 1956).

⁸ At the same time the whole concept of linguistic level can be reformulated much more rigorously.

⁹ Cf. William G. Moulton, Juncture in modern standard German, *Lg.* 23.225 fn. 14 (1947); H. A. Gleason Jr., *An introduction to descriptive linguistics* 66, 175 (New York, 1955); On quite different grounds, grammatical criteria for phonemic analysis have been advocated by Kenneth L. Pike, Grammatical prerequisites to phonemic analysis, *Word* 3.155 (1947), More on grammatical prerequisites, *Word* 8.106-21 (1952). In a review of Harris's *Methods*, *Lg.* 28.507 fn. 8, Murray Fowler erroneously construes Harris's references to the use of meaning criteria in that work as the major cause of his failure to provide a discovery procedure (though Harris's methods in fact make no use of semantic criteria), and uses this charge of circularity as his strongest criticism of distributional analysis. Cf. also Eugene A. Nida, The analysis of grammatical constituents, *Lg.* 24.173 (1948); Hockett, review of Martinet 341.

mar from the data, there can be no question of a compulsory order in which various units must be specified. Nor is there any notion of circularity of definition involved in the specification of empirical elements, since the latter must be sharply distinguished from definitions in the strict sense, which are agreements on how technical terms are to be introduced into a theory. Furthermore, even if one did establish an order of discovery, and phonemic analysis were prior to syntactic analysis, any dependence of the latter upon the former would, ipso facto, imply a corresponding relation in the opposite direction, and it is hard to see just how an injunction against 'mixture of levels' could be formulated rigorously, unless this injunction is taken to mean complete irrelevance of phonemics to syntax.

Conditions on grammars. If, then, a grammar, once constructed, is validated by the application of some explicit criteria of simplicity (say, minimal number of symbols used), is this sufficient to guarantee that it will add something to our knowledge of language or languages? Shall we still have to regard similarities in speech behavior among various languages to be some fortuitous convergence of otherwise unrelated linguistic evolutions? Speakers of every language employ phonemes, morphemes, sentences, immediate constituents, vowels, consonants, and more or less rigid ordering of these in utterances. Speakers of all languages acquire the ability during childhood to extend indefinitely the use of the grammatical patterns of their language to produce an endless stream of new sentences, no one of which may ever have been uttered before in all linguistic history, but each of which is nevertheless immediately recognizable by any of those speakers to be fully grammatical. If a grammar is to be of any more general interest than a handbook for learning the language of which it purports to be the description, then clearly it must satisfy more conditions than maximum simplicity.

To ensure the significance of a grammar, Chomsky would impose two further types of conditions upon it: a set of so-called 'internally linguistic' conditions, and a set of 'external conditions' (§6.1). If a grammar is to explicate the kinds of linguistic behavior mentioned above, it must meet not only the internal requirement of simplicity, but must be so constructed that all the linguistic units and concepts used can be shown to be special cases of more general definitions embodied in a linguistic meta-theory, a theory of language; otherwise there would be no reason to expect that our knowledge of how speakers of any one language used it to communicate had any relation at all to what the grammar of other languages could elucidate about communication among their speakers.

Furthermore, if it is to be useful in explicating the behavior of speakers, a grammar must be able to satisfy certain external, empirical requirements. It must permit us to generate automatically all and only the grammatical sentences of the language, else it could not be called a description at all. Moreover, these sentences cannot be restricted to some finite corpus, say a text, or the set of all sentences which had been uttered up to the time of the construction of the grammar, else it could not account for the fact that speakers are able to extend the corpus indefinitely.

Clearly, some kind of recursive rules will be required in order that a finite grammar generate an infinite set of sentences. Although the morpheme inventory of a language is finite in size at any one time, and sentences must be constructed of

morpheme sequences, the number of sentences is unlimited because there is no meaningful upper bound which may be placed on the length of sentences, even though any given sentence must be finite in length.

Furthermore, speakers exhibit a fairly consistent ability to grade sequences of morphemes by degree of acceptability as utterances, even when they are meaningless (as in nonsense verse, like Lewis Carroll's *Jabberwocky*); therefore a grammar must permit the construction of a scale of grammaticality (Chap. 2, fn. 2; Chap. 5, fn. 2). This scale extends from indisputably bona fide sentences such as *Birds sing*, through various degrees of grammaticality as in *Their black, round, squares of milk don't fit today*, *She would have been being silly then*, or *Whom are you seeming?*, all the way to clearly ungrammatical sequences such as *Mine dispose out umpire the*.¹⁰

Still another external condition which one could impose on a useful grammar, one which Chomsky has found particularly expedient in lending support to a given analysis (Chap. 8), is that it explicate our intuitive understanding of ambiguous sentences by providing two or more different automatic derivations for them. It should also yield different derivations for sentences which, though very similar in apparent outward form, are understood differently. Thus, *He bought stock for me* can be understood in two ways: either it means *He bought stock; the stock was for me*, or it means *He bought stock; he did it for me*, and an adequate grammar of English would, we hope, automatically provide two different analyses for that sentence. This is not simply a question of two different meanings for the preposition *for*, but rather of two different phrase-structures; this becomes apparent when we nominalize the sentence with an *ing*-transformation into two different transforms: *His buying stock for me (was ...)* and *His buying stock (was for me)*.

Similarly, the two sentences *It was proved by Fermi* and *It was proved by induction*, though they appear to be of the same outward form, i.e. have the same constituent structure, are understood differently, and this must be accounted for by an adequate grammar. Again, this is not merely a difference in meaning between *Fermi* and *induction*; the first sentence must be related to *Fermi proved it* (by a passive transformation), while the second is related to *X proved it by induction*.

The information source model. Given these various conditions which might be placed upon a grammar, especially the indispensable requirement that it at least permit the generation of all the grammatical sentences of the language, what is the weakest sort of model for language which we could entertain?

It is fashionable for information theoreticians, communications engineers, and some linguists to seek an explication of linguistic phenomena in some statistical model of language.¹¹ While it would be silly to deny the usefulness of such models

¹⁰ Cf. David L. Olmsted, review of *Psycholinguistics*, *Lg.* 31.50 (1955).

¹¹ See Hockett, *A manual of phonology* 3-14 (Baltimore, 1955). Cf. also Claude E. Shannon and Warren Weaver, *The mathematical theory of communication* (Urbana, Ill., 1949); Hockett, review of Shannon and Weaver 86-7; Benoit Mandelbrot, *Structure formelle des textes et communication*, *Word* 10.1-27 (1954); id., *Simple games of strategy occurring in communication through natural languages*, *Trans. IRE*, PGIT-3. 124-37 (1954).

for an understanding of certain problems in science, such as the behavior of ideal gases, there does seem to be good reason to believe that language, and more particularly grammar, is essentially a nonstatistical structure, and is rather to be understood with the help of combinatory or algebraic models.

One recent suggestion is that speech be considered the output of a kind of finite automaton which generates sentences from left to right as it changes from state to state with its conditional probabilities, a kind of Markov chain of predispositions to emit the next symbol with a likelihood dependent only upon which state the machine is in at the moment. When these states are identified with the immediately preceding n symbols in a string of linguistic symbols, such a model yields its special case, the n th-order approximation to bona fide text, where grammaticalness might be interpreted as 'high-order approximation'.¹² There is at least this one strong objection to all such conceptions of language: there is no reason whatever to believe that the relative frequency of occurrence (or limit thereof, i.e. the probability) of a sentence or any other sequence, even if this could be rigorously determined, has anything at all to do with its grammaticalness. Thus, there are countless sentences so short as to be of undisputed grammaticalness, whose relative frequency of occurrence must be very much smaller than such longer and more involved, but more common, sentences as *Any resemblance to persons living or dead is entirely coincidental*. If linguistic description is merely a matter of replacing all very low probabilities with zero, then many an impeccable sentence such as *Birds eat* would have to be rejected because it is hardly ever said.

Furthermore, the fact that among successive orders of approximation there appear randomly various grammatical sequences as well as various ungrammatical ones at EVERY level of approximation shows that the order is not relevant to grammaticalness.

A somewhat more powerful model of a finite-state machine, the first in order which could be at all seriously considered as a model of language, would permit the generation of infinitely many sentences using a finite apparatus, a so-called finite-state Markov process, as information source. Chomsky has shown¹³ this model to be inadequate to account for natural language with its recursive nesting of constructions within one another. Chapter 3 is a nontechnical review of his arguments.

Involved in all such considerations is the notion that a grammar may be described as though it were a kind of machine, of whatever sort. This is not intended as another banal example of scientism on the part of the modern grammarian, for there is a clear relation between the essential properties of a mechanical device and the structure of a scientific theory. In order that a theory, and therefore also a grammar, be perfectly public and reproducible, but at the same time effective, it is necessary that the predictions afforded by the theory be an automatic consequence of its premises and arguments. If then the appropriate notations be set up, the derivation of its predictions from its premises may be

¹² Shannon and Weaver 13-5; George A. Miller, *Language and communication* 83-6 (New York, 1951); Kellogg Wilson, *Psycholinguistics: A survey of theory and research problems* 46 (Baltimore, 1954).

¹³ Three models (cf. fn. 1).

translated into machine terms, and, depending upon how complicated an algebra was employed in the theory, the derivations may be mechanized inside an actual physical machine, such as an electronic computer. There is a body of mathematical literature dealing in detail with the properties of such machines.¹⁴ An ordinary electronic computer would suffice to generate the sentences permitted by a finite-state model, mentioned above. To yield derivations for the sentences of a natural language, a machine would require more power than this, say at least as much as an automaton with infinite memory (or, as it is usually called, a Turing machine).¹⁵ But there is as yet no indication that any of the theorems from the mathematics of Turing machines is at all revealing linguistically. In other words, it is of no interest that a Turing machine would suffice in power to account for grammar; the important question is: exactly WHICH particular Turing machines out of the infinite possible kinds best explicate the linguistic behavior of speakers. Specification of the internal structure which such a machine would have to have is just the study of grammar, and the algebra which it would have to obey is not likely to be an interpretation of any system known or studied independently in classical mathematics.

Phrase-structure grammar. The most immediately interesting feature of Chomsky's researches into English syntax and grammatical theory is his rigorous attempt to construct a grammar on the basis of a carefully axiomatized and consistently detailed level of 'phrase-structure', that is, roughly, 'bracketing' or, as it is usually termed in linguistics, immediate-constituent analysis (Chap. 4).¹⁶ He has brought out clearly the resulting difficulties which inevitably beset any such serious attempt.

While the communication engineer assumes language to be an interpretation of some weak version of Turing machine, the linguist, more sophisticated in matters of natural language, has assumed that the sentences of a language may each be analyzed into a linearly concatenated sequence of immediate constituents, and that this bracketing or parsing operation may be performed at various levels of generality to yield a hierarchical branching-diagram, such that any unit at any level is just a certain continuous string within some sentence or else a class of such strings drawn from different but grammatically equivalent sentences.

The difficulties which arise in such a simple-minded model, e.g. with discontinuous components and portmanteau morphemes, have been the topic of much recent methodological discussion.¹⁷ But, previous to Chomsky's attempt, no one has really taken seriously any set of criteria of adequacy and simplicity in

¹⁴ E.g. S. C. Kleene, *Introduction to metamathematics*, esp. Chap. 13 (Princeton, 1952).

¹⁵ Ibid.; see also J. G. Kemeny, Man viewed as a machine, *Scientific American*, April 1955, 58-67, for a particularly graphic description of the Turing machine. Even if such a machine were *powerful* enough to reproduce all of human behavior, as Kemeny suggests, it does not seem to this reviewer any more fruitful as an approach to an understanding of this behavior than computer theory is revealing in the explication of grammatical facts.

¹⁶ The independence of phrase-structure from forms, order, and construction has also been pointed out by Hockett, Two models of grammatical description, *Word* 10.218-20 (1954). Wells also notes the independence of sequence from 'construction' (= phrase-structure) in IC analysis.

¹⁷ Cf. Floyd G. Lounsbury, A semantic analysis of the Pawnee kinship usage, *Lg.* 32.159-62 (1956); Sol Saporta, Morph, morpheme, archimorpheme, *Word* 12.9-14 (1956).

the construction of a grammar of any one language, has ever really followed out to its last implication any consistent method of representing the sentences of any one language in a revealing, intuitively satisfying way. Chomsky is, then, one of the first to emphasize clearly the ever compounding difficulties attendant upon any such grammatical description based exclusively on such a phrase-structure model, and he is the first to offer a constructive suggestion for circumventing these difficulties.

Before going on to this suggestion, let us first note some of these inadequacies of a phrase-structure grammar (Chap. 5). First, there are many instances of sentences understood quite differently, but for which there seem to be no grounds within phrase-structure for assigning different representations without introducing either intolerably complex or arbitrarily unmotivated or unintuitive machinery. For example, there is no way in such a grammar to say that *What are you looking for?* and *What are you running for?* have different structures. As is shown by a simple paraphrase, the first sentence contains a prepositional phrase *for what*, while the second contains an interrogative *what for* 'why', but there is no mechanism for associating a sentence with its paraphrase, and the two sentences appear therefore to have identical constituent structures.

Similarly, there are many cases of ambiguous sentences for which only a single analysis seems justified.¹⁸ Thus, the sentence *This teacher's marks are very low* is understood in two ways: 'This teacher gives low marks' or 'This teacher gets low marks'. There is no reason, however, for assigning two different immediate-constituent analyses, nor is there any word in the sentence which may be said to have two different lexical meanings (homophony).

Then there are also cases where two or more analyses would result for a single unambiguous construction, if we agree to apply the most obvious criteria of simplicity to our grammar. For example, within English phrase-structure there would be grounds for assigning to the sentence *The dog is barking* both the analysis Noun phrase (*the dog*) + Auxil. verb phrase (*is -ing*) + Verb phrase (*bark*), and also the analysis Noun phrase (*the dog*) + Copula (*is*) + Adjective (*barking*). The latter analysis is clearly a counterintuitive result, but considerations of simplicity in the analysis of certain other sentences require such a treatment: *The dog is a friendly animal* (NP + *is* + NP) and *Barking is a sign of excitement* (NP + *is* + NP), therefore: NP = V-ing.

Another type of difficulty is encountered when we try to formulate such rules as that of conjunction. To specify which sequences may appear on the two sides of the conjunction *and*, it would be necessary to designate more than the internal immediate-constituent construction of each component, for they must be constituents of the same kind and derivation to yield fully grammatical conjunctions. Furthermore, even if conjunctions could be described simply as constituents of the same internal and external structure, connected by means of a conjunction morpheme, sentences of the following sort could not be described as 'imperative plus declarative': *Hurry up or you'll be late, Come, here and I'll tell you a story* unless the notion of imperative and declarative sentence types is relinquished.

¹⁸ The inadequacy of IC analysis of ambiguous sentences is also recognized by Hockett, *Two models* 218.

This whole notion of sentence type is, as a matter of fact, quite unmotivated in a phrase-structure grammar. Thus, there is no way to show that *John hit Bill* and *Bill was hit by John* are related, though of different constituent structure, in a way that *John hit Bill* and *Bill hit John* are not, though of similar structure.

Another notion which it would seem an adequate grammar should explicate is the fairly consistent choice of one type of construction to be more basic or central than another more marginal or derived type. Thus, active sentences are thought to underlie passives (as in the previous paragraph), statements are more basic than questions, main clauses are more central structures than dependent clauses. But there is no compelling reason on the basis of constituent structure to order constructions in this natural way.

Perhaps the most severe defect of a grammar expressed exclusively in terms of a phrase-structure hierarchy, or branching-diagrams, is the extreme complexity required even for the simplest type of sentences, and the great difficulty of stating this phrase-structure in terms of units which may contain one another as constituents. While a branching-diagram can be constructed individually for any one given sentence (i.e. every sentence has at least one specifiable phrase-structure), there is no set of expansion or parsing rules which will yield properly the phrase-structure of ALL the sentences, unless extensive portions of the grammar are restated several times. If these uneconomical repetitions are permitted, then the grammar fails to state the near identity in structure between those parts which must be repeated. For example, all of the mechanism which provides the proper combinations of adjective and noun in such sentences as NP + *is* + A (*Roses are red*, *Men are numerous*, but not *This man is numerous*) will have to be repeated for such sentences as (The +) A + NP + VP (*Red roses smell good*, *Numerous men go there*, but not *This numerous man is coming*).

Transformations. One of the two most far-reaching results of Chomsky's study is then his discovery of a new level of linguistic structure which at once sweeps away most of the difficulties encountered in any attempt to extend phrase structure beyond a description of the most central declarative sentences of the language. (Chap. 7.) For a given rule to apply within this level of grammatical description the derivational history of an expression must be taken into account for each rule serves to convert, or transform, one constituent structure into another.

The basic idea behind this new level of structure was obviously derived from those manipulations characteristic of Harris's discourse analysis: a sentence whose phrase structure differs from those already set up in the grammar may nevertheless exhibit a whole set of internal correlations or selections identical with those found in simpler, already described sentences.¹⁹ In other words, a more

¹⁹ But it is also interesting to note that, in a sense, transformational analysis is essentially a formalization of a long-accepted, traditional approach to grammatical relations. To cite only a single example of classical grammatical thought which is basically a kind of transformation theory: 'It is different when we come to such a combination as *an early riser*, which it is quite impossible to turn into *a riser who is early*. Here the adjunct is a shifted subjunct of the verb contained in the substantive *riser*: *he rises* (vb) *early* (adv) = *he is an early* (adj) *riser* (sb)'—Otto Jespersen, *A modern English grammar* 2.283 §12.12. Here Jes-

marginal sentence or construction may be described as a transform of some underlying, more central structure. Harris, in order to find some core of essential, basic propositions of which all the various sentences of a discourse might be said to be merely various versions, has permitted different sentences to be collapsed into prototypes if they differ in storable grammatical ways from one another; e.g. a passive sentence may be transformed into its active form if this will facilitate the collapse.²⁰

Chomsky, on the basis of somewhat different motives, has been led to set up a whole level of grammatical transformations to deal with all the difficulties encountered in trying to state explicitly a complete and simple immediate-constituent grammar. Here the phrase-structure rules need be used to generate only a central core of simplest sentences, the *KERNEL*, in which only a very limited number of sentence types appears (roughly, the shortest active indicative assertions). All other grammatical sentences of the language can be generated by means of these transformations and may be said to have a derived constituent structure. Now, for example, *What are you looking for?* can be described as a what-question transform of *You are looking for it*, and *What are you running for?* as a why-question transform from *You are running*. Or we may say that passives are less central than actives since the transformation which generates passives from actives is not reversible, as there is no way of identifying the prepositional phrase with *by* which contains the subject of the active. Thus, *The blow hurt John* → *John was hurt by the blow*, but not *John was hurt by the bridge* → **The bridge hurt John*, or *John was elected by the following week* → **The following week elected John*.

The anatomy of grammar. Grammar is now seen as a structure of three main interrelated levels of rules, each developed as a different kind of concatenation algebra. The lowest level contains morphophonemic and phonemic rules in which there is a relation, for the most part order-preserving, between the parts of each representation of an utterance and the temporally arranged parts of the real utterance, and in which, for each sublevel, all the constituents may be developed in a derivation simultaneously. There is a 'highest' level, the level of phrase-structure, which yields branching diagrams, in which any given rule can develop only one constituent at a time in order that each rule be required to take into consideration only the results of the immediately preceding rule²¹ and in which there are no clearly definable sublevels, but rather only whole sets of representa-

persen correctly perceives that the difference in the way that we understand, on the one hand, *he is an early riser*, and, on the other, *he is an early bird* can be accounted for by regarding the former as a transform of *he rises early*, the latter as a transform of the pair *he is an X* and *early bird* (which in turn is a transform of *the bird is early*).

²⁰ Harris, *Lg.* 28.18-25; Co-occurrence and transformation.

²¹ For example, suppose that the representation of a sentence, NP + VP (i.e. noun-phrase plus verb-phrase), is to be developed next, and is rewritten directly as *Dogs* + *can* + *swim*, then there is no specification of the phrase-structure in the result, since there is no way to tell which of the developed elements originated as NP and which as VP. But if the development proceeds in two steps—S = NP + VP, NP = *dogs*, VP = *can* + *swim*, expanding each phrase separately—the phrase-structure is preserved at each step: S = NP + VP = *dogs* + VP = *dogs* + *can* + *swim*, and we can say that *dogs* is a NP and *can* + *swim* is a VP.

tions for any given derived utterance. In the 'intermediate' transformation level each rule must take into consideration the phrase-structure and transformational history of the representations of utterances to which it applies, the rules of transformation may each apply to more than one string, and may furthermore be reapplied to yield the necessary recursiveness for infinite extension.

In such a grammar there would seem to be no particular place of honor accorded to just those considerations which figure most prominently in a traditional grammatical description, viz. so-called 'morphemics'. Descriptive grammatical sketches are, in large part at least, simply more or less elaborate morpheme-order charts. It is seldom found necessary to justify either the particular classes of morphemes selected for detailed treatment (say, bases which precede final *-s* plural, but not bases which precede final *-y* as in *muddy, icy, tiny*), even when they are quite small (as in adjectives in *-er*), or the particular diagnostic stigma chosen to distinguish the members from all other morphemes. But now, viewing a grammar as a theory which will generate all and only grammatical sentences by means of naturally chosen, maximally simple, unrepeatable rules, all bases and all those affixal morphemes which are required for stating phrase-structures (because they are correlated with other morphemes external to their own base or construction) will appear somewhere among the constituent-structure rules. And after the application of all transformations and phrase-structure expansions, before the application of morphophonemic and lower-level rewritings, each sentence will be represented morphemically; the concatenated units at this level are MORPHEMES, for the most part inflectional morphemes and morphological heads, i.e. bases.

Before morphemes can be spelled out phonemically, it will of course usually be necessary to pass through one or more sublevels of mappings which will select the appropriate morphophonemic shapes of heterogeneous morphemes when these selections can be stated simply in terms of the surrounding morphemic environment. (Here e.g. the various shapes for the past-tense verbal morpheme will be selected according as the base is weak, strong, irregular, etc.)

Halle has pointed out²² the following very interesting consequence of basing the grammar upon rule-simplicity rather than upon arbitrary and unmotivated principles, such as bi-uniqueness of phonemic transcription. In American descriptive grammars, allomorphic selections conditioned by both the choice of the surrounding morphemes and by the more immediate phonemic environment are all lumped together (sometimes in two subdivisions) into a morphophonemics, and then rules for the allophonic selections are stated as part of the phonology, with the usual insistence that the phonemic transcription preserve bi-uniqueness and phonetic reality, i.e. a unique transcription for every different perceived segment. The lack of certain phonemic contrasts in positions of neutralization is either ignored or said to devolve upon the defective distribution of the (neutralized) phonemes. In many cases, however, this division entails needless duplication of statement and results therefore in loss of generality in the grammar.

Consider the common case, as in the following Turkish example, of final

²² Personal communication. Halle will adopt this view in his forthcoming work *The sounds of Russian*.

devoicing, or neutralization of voiced-voiceless contrasts in final positions: final morphophonemic *b, d, ğ, g, r, and l* are devoiced to [p, t, č, k, q] and [r, l̥]. Since the traditional treatment recognizes [r, l̥] (devoiced *r* and *l*) as allophones of /*r, l*/, but [p, t, č, k, q] as separate phonemes, the rule of devoicing must be stated twice (if we would avoid setting up two new *r* and *l* phonemes just for this case), once as a morphophonemic rule for /*b, d, ğ, g, r, g*/ and again as an allophonic rule for /*r, l*/. In Chomsky's (and Halle's) conception of grammar, one need consider, after morphophonemic rewritings of the first kind (morphemically conditioned), only one last linguistic level before stating the phonic values of its units in a series of unrepeatable phonetic rules, among which appears this rule of devoicing for all stops and liquids. Although it is not bi-unique, this last linguistically significant level may be called phonemic, especially since it is the first level at which the units are all directly describable in terms of the phonetic system used, say a distinctive-feature system, boundary-markers, and phonetic rules.

As the following successive representations of a sample Turkish sentence show, any transcription lower than the last would contain obviously subphonemic writings, such as the devoiced allophones of the otherwise homogeneous phoneme /*r*/:

{Ahmed + Sg + General + Book + Sg + Def Obj + Mehmed + Sg + Dative + Give + Aorist + 3rd Sg * } (string of morphemes which results from application of all syntactic rules in phrase-structure and transformation levels)

<ahmed + O + O + | kitab + O + I + | mehmed + O + E + | ver + ır + O * > (string of morphophonemes which results from application of all morphophonemic spelling rules for morphemes)

/ahmed + kitabi + mehmede + verir * / (string of phonemes, each characterizable as some maximally nonredundant set of features, which results from application of phonemic rules)

[ahmed | kitabi | mehmede | verir *] (a phonetic representation which results from application of phonetic rules specifying vowel harmony)

[ahmet | kitabi | mehmede | verir *] (a further phonetic representation after application of phonetic rule of final voicing: [b d ğ g r l] = [p t č k q r̥ l̥])

[a.ɰmeɰt̚'k'it̚'abɰmeɰmede'veɣɰr̥ɰ] (a still further phonetic representation after application of other phonetic rules. Continues until all phonemes have been described in terms of ALL available features, at which point the linguistic description is ended, since any further specification of sounds would introduce only new free variations or subliminal distinctions.)

It is not yet completely clear just what structure is required for the level in which phonemic representations (in the sense intended here) are converted into representations in terms of the minimally contrastive phonetic features which the analyst chooses. Something like the following may suffice: the phonemic representation is first rewritten as a series of feature bundles in which the number of features used to specify each phoneme is minimized by the elimination of all redundant features. Then by means of a series of phonetic rules the redundant features are added to each phoneme in some simplest manner until they have all

become specified to such an exact degree that any further distinctions would serve to introduce only phonetic differences which are always in free variation. At this point the linguistic description is at an end, and further specification of noises is a question of physics or physiology. This whole level is then characterized by the unique use of the features mentioned as its alphabet. The successive representations in a derivation will gradually approach biuniqueness, but it is never necessary to require this condition.

Chomsky avoids those philosophical problems entailed by the use of class notation in the grammar, such as were of concern to early phonemicists: is a phoneme an actual noise-occurrence (a phone-token), a class of similar noise-occurrences (a phone-type), a class of similar phone-types (allophone), a class of allophones, or merely (sic) an abstraction corresponding to one of these, or, say, a certain pair of these, as in the Prague concept of the phoneme as a phonological opposition? Once a grammar has been acknowledged to be the THEORY of some language, all grammatical units which appear in it may be accorded the same status as the notions in any physical or chemical theory. For instance, the volume which appears, symbolized by the letter *V*, in the gas law $pV = nRT$ is not an actual gas volume, nor even a class of gas volumes, it represents volumes in a theory of ideal gas behavior. So too we shall say that a concept such as phoneme, noun-phrase, or auxiliary verb, which occurs in the algebra of a language, in its grammar, REPRESENTS a certain string of noises, or a certain type of string of noises. (Within the phrase-structure part of the grammar, then, any string might have many different representations: e.g. $NP = T + A + N = T + \textit{black} + N = T + A + \textit{shoe} = \textit{the} + A + \textit{shoe} = \textit{the} + \textit{black} + \textit{shoe} = \text{etc.}$) And it is just one of the tasks of linguistic meta-theory to specify exactly the nature of this relation of naming or representing which holds between concepts in the grammar and real pieces of speech, or between one representation of a string and another.

Needless to say, the structure of such algebras of grammatical levels, representation relations, etc. has nothing to do with the particular graphic devices used to mention or manipulate them on paper. A valid grammatical statement is just as valid whether it is affirmed in an abstruse algebraic notation or in plain words. Finally, it is not the use of mathematical symbols and formulaic statement which renders a given treatment 'formal'. Mathematical formality in an interpreted system means rigorous statement of how each symbol is related to the empirical datum which it represents.

Theory of grammar. The other most important result of Chomsky's theory of language is his very strict axiomatization of linguistic theory. He has chosen to take seriously the requirement that a grammar be not merely an arbitrary reorganization of some corpus, but (in a specifiable sense) a simplest machine which will generate all and only the grammatical sentences of a language. But in order to devise an adequate measure of simplicity, it is incontrovertably necessary first to specify in exact and minute detail just what the internal structure of a grammar and a grammatical level is, and just exactly how the levels are related to one another. Then the grammar of any one language can be required to meet a set of relatively weak external requirements of conformity to the meta-theory of grammar.

It may be that, in our great enthusiasm to introduce linguistic relativity as an

antidote to prescriptive classicism, we have also thrown out all grammatical theory, like the proverbial baby with the bath water. The theory of universal grammatical categories, like Ptolemaic astronomy, while quite wrong, was more in need of revision than of repudiation. This is not to say that the descriptive linguist has no theory of language at all. He believes, it is true, that all languages have phonemes and morphemes, that all sentences have immediate constituents, etc. But at the same time, he inhibits the development of a comprehensive theory of language by favoring an exaggerated relativity of arbitrarily chosen form classes.

Replacing erroneous semantic notions with strictly specified 'formal' categories was, to be sure, an advance, especially where those categories correspond well to intuitive notions of structural equivalence. The so-called formal analysis has the one advantage of permitting the analysis of any given sentence in the presence of a correct grammar, by reference only to the rules and to the sentence itself, while the older semantic classifications would require additional knowledge of the meanings.²³ But merely specifying the membership of a class by means of more explicit notation, such as a diagnostic environment or an exact list, does not provide any deeper understanding of the category involved.

The correctness of a grammatical statement is assured not by the arbitrarily chosen differentia, semantic or formal, but by the analyst's correct perception of the underlying phrase-structure or transformational history. For example, to separate nouns from verbs in English is certainly necessary, but to do so by defining them as classes of morphemes which may occur before certain final affixes (rather than certain others) is just as arbitrary and ad hoc as is a separation on the basis of 'means person, place, or thing, etc.' versus 'means action, etc.' In the 'formal' analysis there is no motivated reason for choosing the particular affixes used rather than some others, or for using any other differentiae which happen to specify the members of the desired class. Nor is there ever any serious attempt to catalog ALL POSSIBLE classes on the basis of the following affix, as would be necessary if this were the true analytic criterion, for the analyst knows full well that to do so would yield a very detailed but completely vacuous morpheme-order chart, and not the desired categories of noun and verb. Given a set of meaningful categories or morpheme classes, it is always possible to find or devise some explicit 'formal' property which will serve to specify just which elements belong to the list of members of the various categories. This is just as arbitrary as any semantic classification which accomplishes the same task; and as for explicitness, nothing could be more explicit than a simple listing of the members.

Furthermore, the 'formal' analysis is no more generalizing than the semantic. To include *sheep*, *oxen*, *alumni*, etc. among nouns requires special rules, just as special rules are required to accommodate meanings within the semantic specification other than 'person, place, or thing'. Such ad-hoc rules, whether semantic or formal, require special invention for each individual case, and thus add little to our understanding of grammatical categories. Only when an individual category in some language can be shown to be a special case of a more general notion of

²³ Gleason, *Introduction* 92-5.

'grammatical category' applying within a linguistic meta-theory to all languages, are we justified in seeing this as a real step toward elucidating the behavior of speakers of that language.

Meaning in linguistics. So far we have said nothing at all about the vexed question of semantic criteria in linguistics (Chap. 9). There is a very simple explanation for this neglect: if the term 'meaning' is taken in its ordinary, everyday sense, this notion turns out to be simply irrelevant to grammatical theory and analysis. It is however not at all irrelevant to language study, and it may even be that part of linguistic studies which is of the greatest interest to the majority of our profession. But the study of meaning and its relation to grammar has been woefully confused by the widespread confounding of reference, meaning, synonymy, 'differential meaning', informant response, amount of 'information', significance, grammatical equivalence, truth-preserving equivalence, and mutual substitutability. No single concept involved in linguistic tradition has caused such widespread misunderstanding and entailed such a plethora of polemic as has that of meaning, with the possible exception of the 'phonetic law' of the Jung-grammatiker. The linguist not only is beset, as is any other behavioral scientist, with all the classical philosophical problems inherent in the notion of meaning, but also must now deal with the added difficulty of identification with one or another of the several schools of linguistic philosophy at odds with one another over the question of whether and (if so) how meaning enters into linguistic analysis.

There are two other scholarly disciplines in which serious attempts have been made to clarify problems involved in this area: philosophy and psychology. Now, while philosophy seems to have succeeded fairly well in elucidating the notions involved in the relation of denotation, that is, all that is meant by reference or naming, there seems to have been little progress so far in explicating the concept of meaning. Although many formal systems have been constructed to deal with semantics, even the best of them assume as basic, primitive notions the relation of synonymy and/or the property of significance, and it is just these very concepts which are so unclear. It is advisable in any case to avoid entirely any use of the term 'meaning' (as a noun) and to speak instead exclusively of a relation 'means', or to be even more careful, using just 'is significant' (= 'has a meaning') and 'is synonymous with' (= 'has the same meaning as').

As for the second discipline mentioned above, it does not seem at all unlikely that if a full explication of significance and synonymy (and therefore also of 'meaning') is forthcoming, it will be found in some psychological or psycholinguistic theory. But, at least for the present, any such theory is far beyond current achievements in complexity and power.

Although it seems quite clear, at least to Chomsky and the present reviewer, that semantic criteria of analysis are neither useful in nor indeed pertinent to grammar, nevertheless we might expect that insight into grammatical structure, and the ways in which sentences are understood, would yield directly a better understanding of significance and synonymy. This is especially true if we seek to formulate grammatical statements in such a way as best to explain ambiguities in and centrality of structure, and when we employ for this a kind of transforma-

tion for which lexicosemantic content may very well be an invariant. Thus it would be a great step forward if it could be shown that all or most of what is 'meant' by a sentence is contained in the kernel sentences from which it is derived. And much of the obscurity beclouding the idea of meaning may very well have resulted in large measure from a restrictive overconcern with lexical items and dictionary entries to the exclusion of the sentence, for while the former are seen in many but by no means all cases to participate in a relation of denotation or naming, it may be that it is only the latter unit, the sentence, which is truly significant, i.e. has a 'meaning'.

It might be of value to bring together here several clarifying notions on some of the things that meaning CANNOT be and some of the reasons why meaning in its ordinary usage cannot be pertinent to grammatical analysis.²⁴

MEANING AND REFERENCE. The first confusion which we must dispose of, one in which the reviewer himself has indulged,²⁵ is the suggestion that meaning is nothing more nor less than REFERENCE, the simple, fairly well understood notion of denotation between an expression and a material object which it denotes. However, there are many expressions which refer to the same object but are not synonymous, for example *the largest city in the world* and *London*; both expressions denote the city of London, but the meaning of the former, as well as that of the latter (if indeed it may be said to have any meaning at all) will remain unaffected after New York has surpassed London in size, though they will then have begun to denote two different objects. Synonyms, if they have any denotation at all, are always co-referent; thus synonymy is a narrower, more complex, and obscurer notion than co-reference, and the latter is a necessary but not a sufficient condition for the former, for all meaningful, referential expressions.²⁶

MEANING AND DISTRIBUTION. It has been suggested, especially often by linguists, whose attention has been focused on linguistic form, that meaning is just a certain kind of distribution in linguistic frames.²⁷ This would presumably mean that the two expressions having the same meaning, again in the ordinary sense of meaning, must share a common distribution, i.e. they must be mutually substitutable, holding some parameter constant, and this distributional fact must serve then to explicate synonymy.

Some logicians have proposed that meaning be explicated with the notion of

²⁴ For similar discussions, see Yehoshua Bar-Hillel, *Logical syntax and semantics*, *Lg.* 30.230-7 (1954), and an answer by Chomsky, *Logical syntax and semantics: Their linguistic relevance*, *Lg.* 31.36-45 (1955).

²⁵ Meaning in three linguistic theories, read before the Linguistic Forum, University of Michigan, in 1953, and again before the Washington Linguistic Club in 1954.

²⁶ Note that some expressions do not seem to have any meaning, in its ordinary sense, e.g. *to in he likes to sing*. Other expressions have no reference, e.g. *the average American* and *mermaid*, unless we are willing to people the universe with countless imaginary objects. Cf. also C. E. Bazell, *The choice of criteria in structural linguistics*, *Word* 10.132 (1954); Miller, *Language and communication* 160.

²⁷ See Nida, *A system for the description of semantic elements*, *Word* 7.1-14 (1951); id., *The identification of morphemes*, *Lg.* 24.430 (1948); Norman A. McQuown, review of Harris's *Methods*, *Lg.* 28.501 (1952); and possibly Harris, *Methods* 7 fn. 4, 365 fn. 6, and *Distributional structure* 155-8. But per contra cf. Bar-Hillel's article 233 and Chomsky's reply 44, as well as Miller, *Language and communication* 112.

truth-preserving substitutability. This view is untenable, for in the true sentence *I'm certain that everyone knows quicksilver is quicksilver*, replacement of the last word by its synonym, *mercury*, will render the sentence false.

Mutual substitutability in all linguistic frames with the preservation of grammaticalness is usually taken as the basic distributional criterion of grammatical analysis.²⁸ This notion will not serve as an explication of meaning, however, for it is not a sufficient condition for synonymy, as may be seen from such expressions as *cerise* and *ecru*, which, though they are quite different in meaning, may be freely substituted in all relevant frames used for grammatical analysis with no change in grammaticalness. Nor is this notion a necessary condition for synonymy, for such expressions as *highball* and *a drink of diluted spirits served with ice in a tall glass*, though they are synonymous, will not occur grammatically in the same linguistically diagnostic frames (e.g. the latter does not occur before the plural morpheme or before the noun *glass* as in *highball glass*).

A somewhat weaker proposal—that 'degree of synonymy', if such could indeed be defined, is given by 'degree of similarity of distribution'—is also likely to fail, for there is at least some reason to expect that a word may occur more frequently in the frames of its antonym than in those of its synonyms; thus, if we could specify all the environments in which *dry* occurs, we would probably also find *wet* there more often than (say) *destitute of moisture*.

MEANING AND INFORMANT RESPONSE. Among the empirical bases of grammatical analysis one of the most important is the informant's indication of which utterance fractions are the same and which are different—that is to say, the data which distinguish free variation from contrast.²⁹ While it is almost universally recognized that repetitions of 'sames' must be isolated from contrasting segments or utterances, it has also been almost universally assumed that this information must be semantic. Thus, it is supposed that 'phonemic difference' cannot be recognized before the phonemes have been isolated by the analyst for fear that the contrary would involve circularity. The bland assumption that an assertion of sameness or difference is a semantic criterion has introduced endless confusion into the question of meaning in linguistics and the basis of grammatical analysis (§9.2.6).³⁰

²⁸ It is interesting to note, however, that no author has yet taken such a distributional criterion really seriously and examined ALL possible frames for substitutability, whether to establish synonymy or grammatical equivalence, i.e. form-class membership. What is really used is substitution in 'diagnostic' frames; and these are chosen quite arbitrarily—which is sufficient to show the futility of a discovery procedure for grammatical analysis by substitution techniques.

²⁹ Cf. Gleason, *Introduction* 182.

³⁰ Bloch has attempted to eliminate the need for the informant's response entirely by basing an apparently mechanical analysis scheme upon distributional criteria; see A set of postulates for phonemic analysis, *Lg.* 24.3-46 (1948); Studies in colloquial Japanese IV. Phonemics, *Lg.* 26.89-90 (1950). That his postulate system does not provide a true discovery procedure can be seen at least from its use of nonmechanizable notions such as 'general definition'; see Contrast, *Lg.* 29.59-60 (1953). Lounsbury, while defending complete independence of semantics and linguistics, still assumes that informant response is a semantic criterion; see A semantic analysis of the Pawnee kinship usage, *Lg.* 32.190-1 (1956).

See also Einar Haugen, Directions in modern linguistics, *Lg.* 27.219 (1951); Eli Fischer-Jørgensen, The phonetic basis for identification of phonemic elements, *JASA* 24.611, 615

THE EMPIRICAL BASIS OF LINGUISTIC ANALYSIS. That the informant response test is not a semantic criterion has been clearly demonstrated by Chomsky,³¹ especially for the case of phonemic analysis. He points out with compelling cogency that 'difference of meaning' is completely irrelevant to 'phonemic distinctness', since homonyms prove it to be an insufficient condition, while synonyms prove it to be an unnecessary one.

This point is of such general interest that a brief summary may not be out of place here. The usually stated canon is that if a phonic difference in some environment entails a meaning difference, then the phonic difference is an instance of a phonemic contrast. In other words, difference of meaning is a criterion to distinguish free variation from contrast. Consider the following crucial types, in which the transcriptions followed by an asterisk³² are in question:

- (1) [mi:t* mi: æt ðʌ mi:t* mærkɪt] *Meet me at the meat market.*
- (2) [ðɪs əu:t* ɪz ʌ skvɛə əvɪt*] *This root is a square-root.*
- (3) [ðɪs li:f* ɪz ɪn ouk li:f*] *This leaf is an oak-leaf.*

In case 1 (homophony) the informant will indicate that the two identically transcribed phone-token sequences are different in meaning, yet we know them to be identical phonemically. Therefore, difference in meaning is not a sufficient condition for phonemic contrast.

In case 2 (synonymy) the informant indicates no meaning difference, yet the forms are phonemically distinct. Therefore meaning difference cannot be a necessary condition for contrast. We conclude that semantic contrast is irrelevant to phonemic contrast.

In case 3 (free variation) the lack of meaning difference will lead the analyst to identify [l] and [ɫ] correctly in the given environment.

If it be supposed that meaning difference can nevertheless serve to distinguish case 2 from case 3 for a whole set of phone-token pairs, each illustrating the same proposed phonemic contrast, by the fact that a meaning difference must be found in at least one such pair if the given phonic difference is phonemic, this can only be because the analyst has been able previously to identify corresponding members of the pairs as phonemically identical. For example, in some other illustration of case 2:

- [ðeəz sʊt* ʌn mae su:t*] *There's soot on my suit.*

The informant may indicate that [ʊ] and [u:] are in contrast because of the meaning difference. This can be used to verify the phonemic contrast of /ʊ/ and

(1952); Lounsbury, *Oneida verb morphology* 16 fn. 8 (New Haven, 1953); Paul L. Garvin, review of Jakobson, Fant, and Halle's *Preliminaries to speech analysis*, *Lg.* 29.476 (1953). Harris avoids this confusion in phonemic analysis by advocating the use of pair-tests (*Methods* 31-2), and alludes to a distinction between informant response and semantic criteria in the phrase 'meaning-like distinction between utterances which are not repetitions of each other' (7 fn. 4); see also 29 fn. 1, but 173, 363 (appendix to §12.41), and 365 fn. 6. The confusion is also avoided by Halle, *The strategy of phonemics*, *Word* 10.200 (1954), and by Hockett, *Manual* 144-5.

³¹ Semantic considerations (cf. fn. 1).

³² The starred examples are taken from Eli Fischer-Jørgensen's discussion of this question in her article *The commutation test and its application to phonemic analysis*, *For Roman Jakobson* 140-51 (The Hague, 1956).

/u/ only if the [ʊ] of [sʊt] has been identified with that of [əʊt] and the [u:] of [su:t] with that of [əu:t]. But this cannot be done on the basis of the meaning-difference criterion, since now the environments are no longer identical.

The use of such meaning criteria to distinguish free variation from contrast appears even more hopeless when we note the many cases in which two phonotypes are in contrast in one environment but in free variation in another. This is the case with many pairs of English vocalic nuclei in contrast under strong stress but freely interchangeable under weak stress, such as /ow/, /o/, and possibly /ʌ/; or, for many dialects, /iy/ and /i/ before /g/ or /ŋ/, and /t/ and /ɛ/ before /r/.

The method actually employed by the analyst in all such crucial cases is some version of the 'pair-test'³³ (§9.2.4), with or without collecting the superfluous information on meaning. It is important to keep in mind that such tests are designed not only to determine whether two segments sound alike or different to the native speaker for one presentation, but also to determine whether the identification or distinction is made consistently. For this reason, the tested pair is presented to the informant in massive, randomized replication. Furthermore, by tape resplicing techniques such comparisons as that of the [u:] of [su:t] with the [ʊ] of [əʊt] may be made.

It has been objected that an informant might easily learn to distinguish free variants, as in case 3 above, and thus invalidate the results. This is just the reason for including a consistency test, for if an informant does in fact distinguish [l] from [ɫ] consistently in all examples tested, then the distinction is not a case of free variation at all, but one of synonymy, and the analyst might expect eventually to uncover nonsynonymous examples.

The same arguments for the independence of grammatical intuition from the meanings of the forms may be given for other levels of analysis.³⁴ Difference of meaning cannot be used as evidence for morphemic distinctness without introducing a vast number of otiose morphemic splits. For example, the morpheme *yellow* has several unrelated meanings, such as 'color between orange and green', 'cowardly', 'venal', but there would be no grammatical advantage in morphemically separating any two semantically distinct instances in identical environments. Similarly, meaning contrast is not a sufficient criterion for difference in constituent structure; in the following sentence, for either of its two meanings, the constituent structure is identical morpheme by morpheme: *It was prohibited by a new law*, which means either 'A new law prohibited it' or else '... prohibited it by means of a new law.'

³³ For a clear statement of the pair-test technique in phonemic analysis, see Harris, *Methods* 31-2; id., *Distributional structure*, *Word* 10.158-9 (1954); Halle, *Strategy*, *Word* 10.200 (1954); Chomsky, *Semantic considerations*.

³⁴ It hardly seems necessary to document the use of a semantic criterion in the identification of morphemes, since this principle is almost universal in contemporary linguistic works. See for instance Bloch, *English verb inflection*, *Lg.* 23.399-418 (1947); Hockett, *Problems of morphemic analysis*, *Lg.* 23.341 (1947); Moulton, *Juncture*, *Lg.* 23.218 (1947); Nida, *System*, *Word* 7.2 (1951); Lounsbury, *Oneida verb morphology* 11; C. F. Voegelin, *Distinctive features and meaning equivalence*, *Lg.* 24.133 (1948); John Lotz, *Speech and language*, *JASA* 22.713-4 (1950); Bar-Hillel, *Logical syntax*, *Lg.* 30.230 (1954); Gleason, *Introduction* 54-5, 77, 79, 109.

Meaning difference is also not a necessary criterion either for morphemic distinctness or for constructional contrast. *Nearsighted* and *myopic*, though identical in meaning, are different morphemically. *He took off his coat* is constructionally different from *He took his coat off*, but they are synonymous. If it be objected that, after all, no two tokens are ever really quite synonymous,³⁵ this is tantamount to surrendering semantic distinction as an analytic criterion.

Often the canon is modified somewhat to read: any two phonemically identical tokens which share some common ELEMENT of meaning are morphemically identical.³⁶ This hypothetical common element, if indeed meanings have elements, is the same kind of ad-hoc designator of morpheme membership as the arbitrary differentia mentioned above, a designator of form-class membership. It is just as easy (or difficult) to discover something common to the meanings of *yellow* (color) and *yellow* (cowardly) as to the meanings of the morphemically distinct *bat* (baseball) and *bat* (animal), if the notion of meaning is extended sufficiently. (We consider the two cases of *bat* to be morphemically different because they are derived from the two different grammatical categories animate noun and inanimate noun, as required by our sentence-generating grammar of English. The two instances of *yellow*, on the other hand, are, to the best of our knowledge, categorically identical adjectives.)

Recognizing that, if meaning were explained in terms of a person's total social and psychological response to expressions, there would probably be no true synonyms, some have suggested that grammatical analysis must be based upon some notion of DEGREE of semantic similarity.³⁷ No one, however, has yet shown how meanings might be quantified or how in practice a degree of synonymy could be used to isolate grammatical units.

Many writers have treated meaning in linguistics as though it could be partitioned into various 'kinds of meaning'. In particular they have attempted to isolate among the parts of total meaning that portion determined solely by the linguistic environment, the so-called 'structural' or 'grammatical' meaning, as distinguished from the lexical or connotative meaning.³⁸ This would seem to be nothing more than a strange and unorthodox use of the word 'meaning' to denote linguistic distribution.

Others, seeking to allay feelings of guilt at the use of apparently semantic criteria when testing for distinctness of units, have emphasized a strict separation of 'meaning' from 'differential meaning'.³⁹ This use of a very misleading term, differential meaning, for the major empirical datum of linguistic analysis merely compounds the confusion, for the term is clearly a misnomer either for 'difference

³⁵ See for instance Nida, *System 6*, Identification 431-2.

³⁶ See Nida, *System 9*; Ward H. Goodenough, *Componential analysis and the study of meaning*, *Lg.* 32.207-8 (1956); but Harris, *Distributional structure* 152.

³⁷ Nida, *Identification* 437 fn. 40; also Harris, *Distributional structure* 157.

³⁸ See Bloch, *English verb inflection* 399-400 §1.2; Lounsbury, *Oneida verb morphology* 18, Pawnee kinship usage 189; Martin Joos, *Description of language design*, *JASA* 22.708 (1950); Gleason, *Introduction* 55.

³⁹ Especially Henry Lee Smith Jr. and George L. Trager. See for example Smith, review of Carroll's *The study of language*, *Lg.* 31.61 (1955); *Linguistic science and the teaching of English* 11-12 (Cambridge, Mass., 1956).

in meaning' (not a certain KIND of meaning), which is a useless semantic criterion, or for mere 'difference', a criterion whose kind depends upon how the informant's response has been elicited.

If, then, because of the arguments offered above, semantic criteria, such as 'difference in meaning', have been withdrawn as useful evidence for grammatical structure, there would seem to be nothing in the linguistic data to indicate to the analyst which utterances or utterance fractions must be compared and contrasted to test which tentative structural units. First, this is not a valid argument in favor of semantic criteria any more than it is in favor of any other possible criteria, say chemical, political, or theological. Second, as we see from the practice of linguists, though unfortunately not from their own descriptions of linguistic methodology, the criterion actually used in all crucial cases is either the informant's response in carefully designed pair-tests or other elicitation techniques, so constructed as to be completely indifferent to meanings, or else the linguist's own *Sprachgefühl*⁴⁰ is called upon to provide the correct analysis, after which any ad-hoc rule may be devised to designate the results. But any serious and consistent attempt to use synonymy simply yields the wrong answers.

INFORMATION AND MEANING. It might be well to take note of just one last invalid suggestion for the explication of meaning, which has been offered perhaps in the hope that some powerful mathematical treatment could be borrowed from the communications engineer to clarify an obscure semantic notion.⁴¹

Given a code with which messages may be formulated and transmitted and some process of generating messages such that to each code symbol a probability may be attached, then to every message may be associated a quantity, called 'amount of information', which varies appropriately with the uncertainty of the receiver in his identification of the message transmitted.⁴² The amount of information may then be used to devise various other measures, as for the efficiency of a code or the capacity of a transmission channel.

Perhaps because of the somewhat unfortunate choice of the technical term 'information', which is colloquially very close to the term 'meaning,' it has been supposed that the two concepts may be related. But since any number of different meanings may be attached to a message regardless of its information content, this is seen to be a false identification.

⁴⁰ It is precisely this *Sprachgefühl*, this intuitive notion about linguistic structure, which, together with the sentences of a language, forms the empirical basis of grammatical analysis; and it is precisely the purpose of linguistic science to render explicit and rigorous whatever is vague about these intuitive feelings. This is apparently what Hockett means in his discussion of 'empathy' in phonemic analysis, *Manual* 146-7.

⁴¹ See for instance Joos, review of Locke and Booth's *Machine translation of languages*, *Lg.* 32.294 (1956); but per contra cf. Wilson, *Psycholinguistics* 46.

⁴² There have been many popular treatments of information theory since Shannon's original papers in the *Bell Journal*. Both Shannon and Weaver point out the necessity of distinguishing information from meaning; see *The mathematical theory of communication* 3 and 99. See also E. Colin Cherry, A history of the theory of information, *Proc. Inst. Electr. Engrs.* 98.383 (1951); Wilson, *Psycholinguistics* 46; Hockett, review of Shannon and Weaver, *Lg.* 29.89-90 (1953); Miller, *Language and communication* 41.

Some unsolved problems. Full of insights as it is, the three-level sentence-generating conception of grammar described above, like any useful scientific theory, has given rise to a host of new problems. But one of its most powerful features is the possibility which it provides of giving an exact formulation of a number of difficult theoretical questions.

TECHNICAL DIFFICULTIES. First there are innumerable internal 'technical' problems demanding solution. While it does not seem too difficult to state precisely the assumptions and definitions required to erect a rigorously formed level of phrase-structure for kernel sentences and to describe exactly the notions of grammatical category, constituent, the representing relation holding between any analysis and the utterance, and related grammatical ideas, an exact formulation of the level of grammatical transformations is by no means so simple. Before we can say unequivocally just what a transformation is and just how it enters into a grammar, it will be necessary to formulate rigorously an algebra of transformations.⁴³

In particular, transformations will have to be so formulated that the transform is provided with a constituent structure which is capable of entering another following transformation as argument, since some sentences will be generated by the use of more than one transformation. The constituent structure of a kernel sentence is derived by successive expansions in a branching-diagram or derivational tree in such a manner that at every step, i.e. after the application of every rule, any constituent of the resulting representation can be identified as a derived and expanded instance of some more general grammatical category in its tree. But a transformation takes as argument a whole constituent structure, specified by its derivational tree, and converts it by means of additions and subtractions of elements into a new constituent structure, and the elements of the new structure are no longer direct expansions of previously derived elements. Some other way, then, will have to be found to specify the grammatical status of the transform elements.

Consider, for example, the case of an interrogative passive sentence, derived by means of two successive transformations from an active assertion:

John hit Bill (passive) → *Bill was hit by John*

Bill was hit by John (question) → *Was Bill hit by John?*

For the second or question transformation to apply properly, we must be able to recognize automatically that *Bill* is a subject noun-phrase (NP) of the intermediate sentence. This could be specified by the rule that any transform element derived from a NP in the argument shall be a NP in the transform. But this cannot be generalized into an exhaustive rule because of cases like the following.

Suppose that we have an active, affirmative assertion, the subject of which is a conjunction of two genitives. Each genitive must first be obtained by means of some transformation of a kernel sentence, the two must then be joined by means of the conjunction transformation, and then the conjunction must be inserted

⁴³ A good start has already been made in this direction in Chomsky's *The logical structure of linguistic theory*, Chap. 8 (cf. fn. 1).

into the subject position of a kernel sentence by means of still another transformation. Thus:

$$\begin{array}{l}
 \textit{Britain has an offer (Nom.}_1\text{)} \rightarrow \textit{Britain's offer} \\
 \textit{America refuses (Nom.}_2\text{)} \rightarrow \textit{America's refusal} \\
 \left. \begin{array}{l} \textit{Britain's offer} \\ \textit{America's refusal} \end{array} \right\} (\text{Conj.}) \rightarrow \textit{Britain's offer and America's refusal} \\
 \left. \begin{array}{l} \textit{It caused some surprise} \\ \textit{B's offer and A's refusal} \end{array} \right\} (\text{Subj.}) \rightarrow \textit{Britain's offer and America's} \\
 \hspace{10em} \textit{refusal caused some surprise}
 \end{array}$$

Now for the last transformation to apply correctly it is necessary first to have designated the constituent *Britain's offer and America's refusal* as a noun phrase (NP). We might say that it received this status as NP from the NP status of the two underlying constituents from which it was formed. But now this NP status must itself have been established; in this case, it is not possible to derive the NP status of *Britain's offer* or of *America's refusal* from the underlying constituent from which each was transformed, since these underlying forms themselves are not NPs but sentences. Therefore the identity of the transform as a NP must come from elsewhere.

The only other source for the derived constituent structure of transforms, other than the phrase structure of the representations of which they are the transforms, would be comparison with appropriate kernel sentences. If it were possible to state rigorously that *Britain's offer* HAS THE SAME STRUCTURE AS a kernel sequence like *this offer*, we could automatically derive the NP status of *Britain's offer and America's refusal* from the known NP status of *this offer and this refusal*.

This further step, however, introduces a new problem: how to state the meaning of 'has the same structure as'. This might be accomplished if it were possible to establish some set of fundamental categories, such as NP, VP, A (adjective), in terms of which such a comparison with kernel sentences could be made automatically.

Another internal technical problem which awaits exact solution is entailed by the establishment of a grammatical theory upon the notion of grammaticalness of sentences. It is necessary to establish some scale of grammaticalness along which every utterance will lie and which will correspond well to our intuitive feelings about how sentences are construed, and which, furthermore, will be automatically derivable from the general theory of language or from the grammar of each particular language. This is to say that a grammar must explicate our notion that certain structures, while very bizarre, are nevertheless not completely excluded by the pattern of the language, and that certain ones are less excluded than others and perhaps even by a specifiable degree. This question may very well be related to the establishment of the fundamental categories mentioned above.

When one attempts to state exactly and economically in all their great detail the rules for some real natural language, it becomes quite clear that an immediate-constituent or phrase-structure grammar for all the sentences cannot be given without introducing extremely unnatural repetitions of many ad-hoc rules on the one hand, or on the other restating the entire grammar in a much simpler form

by the use of grammatical transformations. There is as yet, however, no rigorous proof that a phrase-structure grammar of a natural language is inherently impossible. While the great simplicity permitted by the introduction of transformations is sufficient reason to accept this theory of grammar, it is at least intellectually unsatisfying not to have such a proof, especially when a corresponding proof can be supplied so easily for certain very simply constructed symbolic languages having some of the same properties as natural language.

By now the reader has no doubt come to appreciate the magnitude of the technical problems involved. There is certainly no cause for discouragement here, for there are several clear programs for research in this area of grammatical theory from which success can be expected. Nor is any incompleteness in the specification of the algebra of transformations a reason to halt practical grammar writing; on the contrary, it is only by writing out the myriad details of real grammars that such problems can be solved.

RECURSIVE SELECTIONS. In the treatment of any one language there is also a large number of specific analytic problems involved that may be solved in various ways, no one of which is entirely satisfactory. This sometimes implies that there may be more devices for generating sentences in a grammar than have been discovered so far.

As an example, we may note that there does not seem to be any one completely natural way to generate noun-noun compounds in English. Just as in the case of the genitive construction, there are strong selection rules determining which particular combinations of noun with noun can occur. The genitive combinations seem for the most part to be directly relatable to one or more kernel sentence and can be derived easily by means of a small number of transformations, thus:

John has a car → *John's car*

John is safe → *John's safety*

John flies → *John's flight*

(but: *a week's wages*, *Verdi's Aida*, *for John's sake*)

But compound nouns are much more difficult to relate to kernel sentences, except for certain types, such as:

a brush for hair → *a hair brush*

a tax on gasoline → *a gasoline tax*

a blow to the body → *a body blow*

a man in the service → *a service man*

No general rule can be given for these because of the failure of others:

a book for cooking ≠ **a cooking book*

an opinion on politics ≠ **a politics opinion*

a road to success ≠ **a success road*

a face in the window ≠ **a window face*

Furthermore, we must take account of the fact that compounding is indefinitely recursive and is productive:

jet engine replacement parts depot =

$[(jet + engine) + (replacement + parts)] + depot$

Suez Canal crisis

Therefore compounds cannot simply be entered in the lexicon. But it is also not convenient to permit this type of recursion to occur in the kernel-generating

phrase-structure, especially since all but the very weakest sort of recursions seem to be best handled as transformations. On the other hand, if they are generated elsewhere, all the selections between the members of the compound will have to be stated twice, once for the kernel sentences containing corresponding $N + \text{Prep} + N$, and again for the compound $N + N$.

Exactly the same problem is encountered also in Turkish. It is compounded there by the fact that there is no convenient source from which to derive not only compound nouns but also genitives. In Turkish, simple affirmative statements corresponding to English sentences with the verb *have* are themselves construed with a genitive. Furthermore, in both languages genitive constructions are recursive like compounds and are therefore best introduced by a transformation.

It does not seem likely that this type of difficulty is confined to English and Turkish. If a satisfactory solution cannot be found within the type of grammatical framework described by Chomsky, there are two alternatives: either there is still another device in language, not yet described, which contributes new structures in the derivation of sentences, or else language is not quite so neatly compartmentalized as it would seem to be at first glance, and perhaps the kernel-generating phrase-structure level is in fact strongly recursive.

PHONEMIC GRAMMAR. In addition to certain algebraic problems there are also many difficulties with the comprehensiveness of the theory of language involved. No attempt has been made here to solve all the different kinds of problems inherent in the study of language; a number of what may be truly grammatical problems are left almost untouched by this treatment. For example, we might demand of a useful grammatical theory that it explicate the following apparently universal behavior of speakers of languages: even in the presence of noise, a native speaker is able to identify correctly with a high degree of success single monosyllables spoken by a second native speaker. One can hardly call upon the grammatical redundancy of the language to explain this behavior; clearly there is something in the sounds themselves which the hearer can identify and classify properly, and this something must vary widely from language to language, since, in general, a person cannot perform very well in such an experiment if he listens to a foreign language.

Since the kind of rules required by the sentence-generating machine that Chomsky describes may be so written that morphemic representations of utterances may be converted quite simply into phonemic and then phonic sequences, the conditions imposed upon such a grammar never entail any description of this sound structure adequate to account for the phonetic behavior mentioned. Nor is there any motivation in such a grammar for describing constraints on the phonemic constituency of syllables or clusters, although all the machinery necessary for a description of phonemic immediate-constituent structure is available. Also nothing is said in such a grammar about the validation of any proposed phonetic framework in terms of which phonemic systems may be analyzed, such as a system of distinctive features. It seems quite possible that phonemic systems, syllable structure, phoneme syntax (sometimes called 'phonotactics'), and other lowest-level constraints may be describable only in terms of an independent 'phoneme grammar', and that the primacy of the phonemic units is just the fact that the phonemic and syntactic grammars, the

former 'looking up' from the phonetic substance, the latter 'looking down' from phrase-structure, adjoin at this level and share precisely these units.

SUPRASEGMENTALS. One currently confused issue about which this conception of grammar has as yet little to say is the question of the suprasegmental phonic elements in the sentence, the stresses, pitches, junctures, features of vowel harmony, etc. There is little doubt but that many suprasegmental features will have to be incorporated directly into the syntax (rather than, say, into the proposed 'phonemic grammar'), and since they tend to have rather extended scope of application over many sentence elements, it is not clear as yet just how they may be introduced into a grammatical description without the construction of some new kind of rules. It seems reasonable, for example, to treat a sentence intonation in English as a kind of morpheme selected in the development of a sentence at a high level in the grammar, in fact, a morpheme which may even undergo the same grammatical transformations as the segmental morphemes. Thus, one might introduce question intonation in the same transformation which inverts subject and verb for yes-no questions, and then, since question-word questions may be produced by a further transformation of yes-no questions, this latter transformation would then merely re-invert the intonation to yield the assertion contour used with this second type of question. In Turkish, on the other hand, the yes-no type of question employs a special affixed morpheme and normal assertion intonation, while the transformation used to introduce an interrogative could also yield the question intonation employed with this kind of question.

ANALYSIS AND SYNTHESIS. It must be emphasized also that while this type of grammar has been constructed to permit the automatic generation of all sentences, there is, of course, no provision for correctly analyzing any given utterance in the presence of the grammar. This might not seem a serious drawback if it were not for the fact that an adequate linguistic theory ought to explain somehow the ability of the hearer to understand utterances as they are received in context, without assuming an impossibly lengthy process of trial and error.⁴⁴ Just as the enormously involved structural complexity and interrelatedness of a natural language seldom leave the analyst in doubt when he is confronted with apparently alternative solutions, so too the convergence of numerous interconnected redundancies in the chain may force unique identifications of structure on the native hearer.

MORPHEMICS. Another difficulty, similar to that mentioned in the section on phonemic grammar, is the lack of motivation in such a grammar to provide a detailed description of much of derivative morphology. That part of the morphological structure of a language which is directly involved in the syntax (that is, largely, the inflectional morphology and bases, or morphological heads), will all be introduced in the course of generating the sentences. But once the sentence has been synthesized as a sequence of such morphemes, there is no natural reason to

⁴⁴ This is not the same question as that of a discovery procedure, which concerns the automatic generation of grammars. This is the weaker notion of automatic generation of the constituent structure of a given sentence, assuming the grammar.

describe the internal structure of words before mapping the morphemes directly into phonemic sequences. It may be that a more adequate specification of the measures of simplicity to be used in judging the excellence of grammars would require an incorporated analysis of words, and therefore of derivational morphemes not yet introduced in the syntax. For example, in English we might permit the derivation of adjectives from nouns within phrase structure, since so many nouns can be correlated with an adjective (*child-ish*, *mudd-y*, *fam-ous*, *use-ful*, *wood-en*, *glob-al*, *telegraph-ic*, etc.), but there does not seem to be much motivation for generating verbs, such as *per-ceive*, *con-ceive*, *re-ceive*, *de-ceive*, etc., from 'morphemes' *re-*, *per-*, etc., even though they are correlated with nouns in *-ception*, adjectives in *-ceptive*, and severe constraints on phoneme sequences within words.

VALIDATION. Finally, we must note that the entire theory of grammar proposed in Chomsky's *Syntactic structures* has been used in the description of only portions of very few languages, though work on discourse analysis has indicated at least tentatively that grammatical transformations will be equally pertinent to a description of many widely varying linguistic structures. Several very limited problems in German, Turkish, and Winnebago have been stated in these terms, and Chomsky himself has worked on Hebrew grammar. But to date the only extensive research on tri-level sentence-generating machines has been with English. Much more cross-lingual validation will be required.

Another opportunity for validation is provided by the rather strict ordering of rules required by such grammars and the concomitant implication of a scale of centrality or generality among syntactic structures. Thus, it is shown that interrogative-word questions must be derived from yes-no questions in English, but not vice versa. If now we should find that English-speaking children generally learn to use the former only after having mastered the latter type, this grammatical fact will provide an explication for the ordering in the learning process. But if children tend to learn in the opposite or in random direction, the ordering among the structures in the grammar will be rather difficult to explain.

Wider implications. In conclusion we take note briefly of several further implications of Chomsky's theory of grammar.

COMMON KERNEL. First of all, this theory permits the definition of a special set of sentences for every language—namely, those sentences derivable by the largely nonrecursive phrase-structure level of the grammar, the kernel sentences. All other sentences, derived by the application of grammatical transformations to kernel sentences, may be considered to be more complex; and it may very well be that the kernel always contains only simple, active, declarative, indicative statements. Just as two disparate texts, after having been subjected to discourse analysis, may be found to be quite similar or even identical in their kernel sentences, so too, many languages, though superficially divergent, may prove to be very similar if compared in their kernel sentences only. In fact, such a comparison may help in the explication of the notion of genetic family; for now, with the help of a more articulated grammar, languages may be compared structurally in the algebras which appear in their grammars.

MACHINE TRANSLATION. Another quite different area in which Chomsky's conception of grammar may prove to be of the utmost importance is the field of machine translation. We speak now no longer of machines as algorithmic schemata, but as concrete mechanical devices, in particular, computers which will automatically translate texts written in one language into corresponding texts in another. It may prove possible, by validating some theory of grammar, to demonstrate that machine translation—or more precisely, an exact solution to the problem of machine translation—is inherently impossible.⁴⁵ On the other hand, it may be quite feasible to construct a best approximation to a perfect translator by building separately three different kinds of device, one corresponding to each of the three levels of linguistic structure, once the algebra of the three levels is clearly understood. In this way we may circumvent the discouraging problems involved in any scheme which attempts merely to render more sophisticated an essentially word-for-word type of translation system.

LEARNING THEORY. Perhaps the most baffling and certainly in the long run by far the most interesting implications of Chomsky's theories will be found in their cohesions with the field of human psychology. Being totally incompetent in this area, I shall allude to only one possible consideration, but one which I find extremely intriguing. If this theory of grammar which we have been discussing can be validated without fundamental changes, then the mechanism which we must attribute to human beings to account for their speech behavior has all the characteristics of a sophisticated scientific theory. We cannot look into a human speaker's head to see just what kind of device he uses there with which to generate the sentences of his language, and so, in the manner of any physical scientist confronted with observations on the world, we can only construct a model which has all the desired properties, that is, which also generates those sentences in the same way as the human speaker. If the model has been rendered maximally general, it should predict correctly the human speaker's future linguistic behavior. We may then attribute the structure of this model to the device in the human head, and say that we understand human speech behavior better than before.

Now it might be objected that the particular structure of our model was entailed, at least in part, by the imposition of a simplicity criterion, and there is no reason to believe that the human speaker is subject to that restriction. We must keep in mind, however, that simplicity in this context is intended in the same sense as in connection with any scientific theory. If we were omniscient and could predict correctly every future event, science would consist merely of a list of those events. In the absence of such universal knowledge we must construct models for the behavior of the world which will account exactly for all data known to date, and then, by rendering them maximally general, hope that they will also predict correctly the observations yet to be made. This is done by choosing among all the alternative models which will account for a given set of known data that model which is maximally simple and elegant—that is, general. We may then attribute the structure of our model to the world itself, until some new and unpredicted datum forces us to change our model—that is, to change our con-

⁴⁵ It already seems quite likely that no finite automaton, such as an electronic computer, will be adequate to generate the nested and recursive properties of natural language.

ception of how the world is really constructed. Granting that this so-called scientific method is valid, it is not too much to assume that human beings talk in the same way that our grammar 'talks', provided the grammar has been constructed as an adequate and maximally general model for that speech behavior.

While admitting that all of the physical data necessary for the construction of a scientific theory must somehow be contained in the observations available to the scientist, we must not suppose that a theory can be constructed by means of a simple process of induction from the data.⁴⁶ A zoologist, observing that each individual specimen of lizard contains a three-chambered heart, may by simple induction generalize these observations to the universal rule that all lizards are characterized by having a three-chambered heart; each specimen contains in its directly observable physical structure an instance of the general rule. But the construction of a scientific theory like the electron theory of chemical valence cannot be accounted for by any such simple inductive reasoning. One cannot observe individual instances of electrons in atoms from which to generalize atomic structure. In the construction of a theory very abstract concepts and models must be postulated and then verified against the data in question. We say that atoms have such and such electrons only because in this way it is possible to account simply and correctly for the observed valences. And when these postulated entities turn out later also to account for many other kinds of observations, such as the photoelectric effect, we are that much more sure that our model does reflect reality. Finally, the more complex the interrelationships among the elements of the model must be in order to have the desired properties, the more unlikely it becomes that an alternative model, yet to be constructed, will ever excel it.

Now the grammar of a natural language has precisely this character. It is a postulated structure containing highly abstract concepts used to account for a speaker's generation of the grammatical sentences of his language. It contains many entities which are not directly observable in the physical structure of any one sentence but which must be hypothesized as a source from which many different sentences may be derived.

For example, there is nothing in the directly observable structure of interrogative sentences that would associate yes-no questions with *wh*-questions. Rather, the latter are more clearly seen to be variants of relative clauses. The simplest way to connect them all up is to assume that yes-no questions are derived from assertions by means of a question transformation which reverses the order of the subject NP and the finite verb. The relative clause is then derived from those same assertions by means of another, a *wh*-transformation, which also reverses order, but only for a noninitial NP, and then adds *wh*- to the NP. Finally, the *wh*-questions are derived by means of this same *wh*-transformation, but this time from yes-no questions, rather than from assertions. These transformations apply BEFORE those which append the affix to the finite verb (the affix is first introduced in the same position before the finite verb as the auxiliary verbs) and which introduce *do* to carry any unattached affix, and the morpho-

⁴⁶ Hempel 23-4, 32-9 (cf. fn. 6 above).

phonemic rules which produce *who* and *what* from *wh-* plus NP_{animate} and NP_{inanimate}, etc.

Assertion: NP₁ + Affix + VP + NP₂ (affix) → NP₁ + VP + Affix + NP₂ (morphoph.) → *John + saw + Bill* (.)

Yes-no question: NP₁ + Affix + VP + NP₂ (question) → Affix + NP₁ + VP + NP₂ (affix) → *do + Affix + NP₁ + VP + NP₂ (morphoph.)* → *Did + John + see + Bill* (?)

wh-relative on subject: NP₁ + Affix + VP + NP₂ (*wh-*) → NP₁ + Affix + VP + NP₂ → *wh + NP₁ + Affix + VP + NP₂ (affix)* → *wh + NP₁ + VP + Affix + NP₂ (morphoph.)* → *who + saw + Bill* (.)

wh-relative on object: NP₁ + Affix + VP + NP₂ (*wh-*) → NP₂ + NP₁ + Af + VP → *wh + NP₂ + NP₁ + Affix + VP (affix)* → *wh + NP₂ + NP₁ + VP + Affix (morphoph.)* → *whom + John + saw* (.)

wh-question on subject: Affix + NP₁ + VP + NP₂ (*wh-*) → NP₁ + Affix + VP + NP₂ → *wh + NP₁ + Affix + VP + NP₂ (affix)* → *wh + NP₁ + VP + Affix + NP₂ (morphoph.)* → *Who + saw + Bill* (?)

wh-question on object: Affix + NP₁ + VP + NP₂ (*wh-*) → NP₂ + Affix + NP₁ + VP → *wh + NP₂ + Affix + NP₁ + VP (affix)* → *wh + NP₂ + do + Affix + NP₁ + VP (morphoph.)* → *Whom + did + John + see* (?)

This derivation shows how *wh*-questions receive their 'interrogativeness' from yes-no questions, of which they are transforms. No amount of simple generalization or induction could have yielded such an elegant result.

We come now to the point of this lengthy discussion of induction versus theory construction. Though it is possible, it is certainly not an easy task for a psychologist to explain the mechanism by means of which a child, confronted with a vast and perplexing array of different stimuli, manages to learn certain things which can be generalized by induction from repeated occurrences. We would not ordinarily suppose that young children are capable of constructing scientific theories. Yet in the case of this typically human and culturally universal phenomenon of speech, the simplest model that we can construct to account for it reveals that a grammar is of the same order as a predictive theory. If we are to account adequately for the indubitable fact that a child by the age of five or six has somehow reconstructed for himself the theory of his language, it would seem that our notions of human learning are due for some considerable sophistication.

Fundamentals of language. By ROMAN JAKOBSON and MORRIS HALLE. (Janua linguarum: Studia memoriae Nicolai van Wijk dedicata, No. 1.) Pp. x, 87. 's-Gravenhage: Mouton & Co., 1956.

Reviewed by MARTIN JOOS, *University of Wisconsin*

This slender volume has a first part signed by both authors on Phonology and Phonetics, and a smaller second part signed only by Jakobson on Two Aspects of Language and Two Types of Aphasic Disorders. I will speak first of the second part. It is a fascinating essay, as we should expect; but here I need to speak of it principally to show how it reveals a dominant trait of Jakobson's way of thinking, or at least of presenting his thought.

It appears possible to describe most cases of aphasia economically by specifying the loss, each in its own degree, of just two normal functions: the referential and the syntactic. In case only the referential function is noticeably lost, the