

NOAM CHOMSKY

Studies on Semantics in Generative Grammar

MOUTON

Minor

JANUA LINGUARUM

STUDIA MEMORIAE
NICOLAI VAN WIJK DEDICATA

edenda curat

C. H. VAN SCHOONEVELD

Indiana University

Series Minor, 107

STUDIES ON SEMANTICS
IN
GENERATIVE GRAMMAR

by

NOAM CHOMSKY

MOUTON PUBLISHERS · THE HAGUE · PARIS · NEW YORK

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 74-189711

First edition: 1972
Second printing: 1975
Third printing: 1980

ISBN 90 279 7964 2

© Copyright 1972 by Mouton Publishers, The Hague, The Netherlands
No part of this book may be translated or reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publishers

Printed in the Netherlands



PREFACE

The three essays that follow take as their point of departure the formulation of grammatical theory presented in such work as J. J. Katz and P. M. Postal, *An Integrated Theory of Linguistic Descriptions*, 1964, and Chomsky, *Aspects of the Theory of Syntax*, 1965. For ease of exposition, I refer to this formulation as the "standard theory". The essays deal with problems that arise within this framework, and present a revision of the standard theory to an "extended standard theory" (EST). The status of deep structure is a central concern in all three essays. In the first, "Remarks on nominalization", the device of syntactic features is exploited to formulate the "lexicalist hypothesis" with regard to derived nominals, and evidence is presented indicating that this hypothesis is correct and that the properties of these structures can be appropriately expressed only in terms of the abstract concept "deep structure", in the sense of the standard theory and EST. The status of deep structure is discussed again in the third essay, where further evidence is presented leading again to the conclusion that a level of deep structure (in the sense of the standard theory and EST) must be postulated. The second essay is concerned with inadequacies of the standard theory, and a more refined theory of semantic interpretation is proposed, giving EST: the grammatical relations of the deep structure remain fundamental for semantic interpretation, determining what have been called "thematic relations" or "case relations", but other aspects of meaning are determined by surface structure. The third essay develops EST further. Both the second and third essays compare EST with alternatives, in particular with the approach now often called "generative semantics", and present

evidence that in the areas of difference, EST is to be preferred on methodological as well as empirical grounds. It is also argued that the areas of difference are more slight than much current work suggests and that many issues that appear to be significant in reality reduce to matters of terminology and notation, when clarified.

N. C.

ACKNOWLEDGEMENTS

Permission to reprint these articles from the following sources is gratefully acknowledged:

"Remarks on Nominalization" from *Readings in English Transformational Grammar*, Roderick A. Jacobs and Peter S. Rosenbaum (eds.) (Waltham, Mass.: Ginn and Co., 1970).

"Deep Structure, Surface Structure, and Semantic Interpretation" from *Studies in General and Oriental Linguistics Presented to Shiro Hattori on the Occasion of His Sixtieth Birthday*, Roman Jakobson and Shigeo Kawamoto (eds.) (Tokyo: TEC Co. Ltd., 1970).

"Some Empirical Issues in the Theory of Transformational Grammar" from *Goals of Linguistic Theory*, Stanley Peters (ed.) (Englewood Cliffs: Prentice-Hall Inc.).

CONTENTS

Preface	5
Acknowledgements	7
Remarks on Nominalization	11
Deep Structure, Surface Structure, and Semantic Interpretation	62
Some Empirical Issues in the Theory of Transformational Grammar	120
Index of Names	203
Subject Index	205

REMARKS ON NOMINALIZATION*

For the purposes of this paper, I will assume without question a certain framework of principles and will explore some of the problems that arise when they are applied in the study of a central area of the syntax of English, and, presumably, any human language.¹

A person who has learned a language has acquired a system of rules that relate sound and meaning in a certain specific way. He has, in other words, acquired a certain competence that he puts to use in producing and understanding speech. The central task of descriptive linguistics is to construct grammars of specific languages, each of which seeks to characterize in a precise way the competence that has been acquired by a speaker of this language. The theory of grammar attempts to discover the formal conditions that must be satisfied by a system of rules that qualifies as the grammar of a human language, the principles that govern the empirical interpretation of such a system, and the factors that determine the selection of a system of the appropriate form on the basis of the data available to the language learner. Such a "universal grammar" (to modify slightly a traditional usage) prescribes a schema that defines implicitly the infinite class of "attainable grammars"; it formulates principles that determine how each such system relates sound and meaning; it provides a procedure of evaluation for grammars of the appropriate form. Abstractly, and under a radical but quite useful idealization, we may then think of language-learning as the process of selecting a grammar of the

* This work was supported in part by the U. S. Air Force [ESD Contract AF19(628)-2487] and the National Institutes of Health (Grant MH-13390-01).

¹ The presupposed framework is discussed in greater detail in a number of recent publications, specifically, J. Katz and P. Postal (1964); Chomsky (1965); and references cited there. For bibliographic references, see pp. 117-119.

appropriate form that relates sound and meaning in a way consistent with the available data and that is valued as highly, in terms of the evaluation measure, as any grammar meeting these empirical conditions.

I will assume that a grammar contains a base consisting of a categorial component (which I will assume to be a context-free grammar) and a lexicon. The lexicon consists of lexical entries, each of which is a system of specified features. The nonterminal vocabulary of the context-free grammar is drawn from a universal and rather limited vocabulary, some aspects of which will be considered below. The context-free grammar generates phrase-markers, with a dummy symbol as one of the terminal elements. A general principle of lexical insertion permits lexical entries to replace the dummy symbol in ways determined by their feature content. The formal object constructed in this way is a DEEP STRUCTURE. The grammar contains a system of transformations, each of which maps phrase-markers into phrase-markers. Application of a sequence of transformations to a deep structure, in accordance with certain universal conditions and certain particular constraints of the grammar in question, determines ultimately a phrase-marker which we call a SURFACE STRUCTURE. The base and the transformational rules constitute the syntax. The grammar contains phonological rules that assign to each surface structure a phonetic representation in a universal phonetic alphabet. Furthermore, it contains semantic rules that assign to each paired deep and surface structure generated by the syntax a semantic interpretation, presumably, in a universal semantics, concerning which little is known in any detail. I will assume, furthermore, that grammatical relations are defined in a general way in terms of configurations within phrase-markers and that semantic interpretation involves only those grammatical relations specified in deep structures (although it may also involve certain properties of surface structures). I will be concerned here with problems of syntax primarily it is clear, however, that phonetic and semantic considerations provide empirical conditions of adequacy that must be met by the syntactic rules.

As anyone who has studied grammatical structures in detail is well aware, a grammar is a tightly organized system; a modification of one part generally involves widespread modifications of other facets. I will make various tacit assumptions about the grammar of English, holding certain parts constant and dealing with questions that arise with regard to properties of other parts of the grammar.

In general, it is to be expected that enrichment of one component of the grammar will permit simplification in other parts. Thus certain descriptive problems can be handled by enriching the lexicon and simplifying the categorial component of the base, or conversely; or by simplifying the base at the cost of greater complexity of transformations, or conversely. The proper balance between various components of the grammar is entirely an empirical issue. We have no a priori insight into the "trading relation" between the various parts. There are no general considerations that settle this matter. In particular, it is senseless to look to the evaluation procedure for the correct answer. Rather, the evaluation procedure must itself be selected on empirical grounds so as to provide whatever answer it is that is correct. It would be pure dogmatism to maintain, without empirical evidence, that the categorial component, or the lexicon, or the transformational component must be narrowly constrained by universal conditions, the variety and complexity of language being attributed to the other components.

Crucial evidence is not easy to obtain, but there can be no doubt as to the empirical nature of the issue. Furthermore, it is often possible to obtain evidence that is relevant to the correct choice of an evaluation measure and hence, indirectly, to the correct decision as to the variety and complexity that universal grammar permits in the several components of the grammar.²

To illustrate the problem in an artificially isolated case, consider

² Needless to say, any specific bit of evidence must be interpreted within a fixed framework of assumptions, themselves subject to question. But in this respect the study of language is no different from any other empirical investigation.

such words as *feel*, which, in surface structure, take predicate phrases as complements. Thus we have such sentences as:

- (1) John felt angry (sad, weak, courageous, above such things, inclined to agree to their request, sorry for what he did, etc.).

We might introduce such expressions into English grammar in various ways. We might extend the categorial component of the base, permitting structures of the form noun phrase-verb-predicate, and specifying *feel* in the lexicon as an item that can appear in prepredicate position in deep structures. Alternatively, we might exclude such structures from the base, and take the deep structures to be of the form noun phrase-verb-sentence, where the underlying structure *John felt* [_S*John be sad*]_S³ is converted to *John felt sad* by a series of transformations. Restricting ourselves to these alternatives for the sake of the illustrative example, we see that one approach extends the base, treating *John felt angry* as a NP-V-Pred expression roughly analogous to *his hair turned gray* or *John felt anger* (NP-V-NP), while the second approach extends the transformational component, treating *John felt angry* as a NP-V-S expression roughly analogous to *John believed that he would win* or *John felt that he was angry*. A priori considerations give us no insight into which of these approaches is correct. There is, in particular, no a priori concept of "evaluation" that informs us whether it is "simpler", in an absolute sense, to complicate the base or the transformational component.

There is, however, relevant empirical evidence, namely, regarding the semantic interpretation of these sentences.⁴ To feel angry is not necessarily to feel that one is angry or to feel oneself to be angry; the same is true of most of the other predicate expressions that appear in such sentences as (1). If we are correct in assuming that it is the grammatical relations of the deep structure that determine the semantic interpretation, it follows that the deep structure of (1) must not be of the NP-V-S form, and that, in fact,

³ Henceforth I shall use labeled brackets to indicate structures in phrase-markers; an expression of the form $X[_A Y]_A Z$ signifies that the string Y is assigned to the category A in the string XYZ .

⁴ There are a number of suggestive remarks on this matter in Kenny (1963).

the correct solution is to extend the base. Some supporting evidence from syntax is that many sentences of the form (1) appear with the progressive aspect (*John is feeling angry*, like *John is feeling anger*, etc.), but the corresponding sentences of the form NP-V-S do not (**John is feeling that he is angry*). This small amount of syntactic and semantic evidence therefore suggests that the evaluation procedure must be selected in such a way as to prefer an elaboration of the base to an elaboration of the transformational component in such a case as this. Of course this empirical hypothesis is extremely strong; the evaluation procedure is a part of universal grammar, and when made precise, the proposal of the preceding sentence will have large-scale effects in the grammars of all languages, effects which must be tested against the empirical evidence exactly as in the single case just cited.

This paper will be devoted to another example of the same general sort, one that is much more crucial for the study of English structure and of linguistic theory as a whole.

Among the various types of nominal expressions in English there are two of particular importance, each roughly of propositional form. Thus corresponding to the sentences of (2) we have the gerundive nominals of (3) and the derived nominals of (4):⁵

- (2) a. John is eager to please.
b. John has refused the offer.
c. John criticized the book.
- (3) a. John's being eager to please
b. John's refusing the offer
c. John's criticizing the book
- (4) a. John's eagerness to please
b. John's refusal of the offer
c. John's criticism of the book

Many differences have been noted between these two types of nominalization. The most striking differences have to do with the

⁵ The fullest discussion of this and related topics is in Lees (1960), from which I will draw freely.

productivity of the process in question, the generality of the relation between the nominal and the associated proposition, and the internal structure of the nominal phrase.

Gerundive nominals can be formed fairly freely from propositions of subject-predicate form, and the relation of meaning between the nominal and the proposition is quite regular. Furthermore, the nominal does not have the internal structure of a noun phrase; thus we cannot replace *John's* by any determiner (e.g., *that*, *the*) in (3), nor can we insert adjectives into the gerundive nominal. These are precisely the consequences that follow, without elaboration or qualifications, from the assumption that gerundive nominalization involves a grammatical transformation from an underlying sentencelike structure. We might assume that one of the forms of NP introduced by rules of the categorial component of the base is (5), and that general rules of affix placement give the freely generated surface forms of the gerundive nominal:⁶

(5) [sNP *nom* (Aspect) VP]_s

The semantic interpretation of a gerundive nominalization is straightforward in terms of the grammatical relations of the underlying proposition in the deep structure.

Derived nominals such as (4) are very different in all of these respects. Productivity is much more restricted, the semantic relations between the associated proposition and the derived nominal are quite varied and idiosyncratic, and the nominal has the internal structure of a noun phrase. I will comment on these matters directly. They raise the question of whether the derived nominals are, in fact, transformationally related to the associated proposi-

⁶ I follow here the proposal in Chomsky (1965, p. 222) that the base rules give structures of the form NP-Aux-VP, with Aux analyzed as Aux₁ (Aspect), Aux₁ being further analyzed as either Tense (Modal) or as various nominalization elements and Aspect as (perfect) (progressive). Forms such as * *John's being reading the book* (but not *John's having been reading the book*) are blocked by a restriction against certain -ing -ing sequences (compare * *John's stopping reading*, *John's having stopped reading*, etc.). Tense and Modal are thus excluded from the gerundive nominal, but not Aspect. Nothing that follows depends on the exact form of the rules for gerundive nominalization, but I think that a good case can be made for this analysis.