

### 3.

## Methodological Preliminaries

### §1. GENERATIVE GRAMMARS AS THEORIES OF LINGUISTIC COMPETENCE

This study will touch on a variety of topics in syntactic theory and English syntax, a few in some detail, several quite superficially, and none exhaustively. It will be concerned with the syntactic component of a generative grammar, that is, with the rules that specify the well-formed strings of minimal syntactically functioning units (*formatives*) and assign structural information of various kinds both to these strings and to strings that deviate from well-formedness in certain respects.

The general framework within which this investigation will proceed has been presented in many places, and some familiarity with the theoretical and descriptive studies listed in the bibliography is presupposed. In this chapter, I shall survey briefly some of the main background assumptions, making no serious attempt here to justify them but only to sketch them clearly.

Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance. This seems to me to have been the position of the founders

---

This chapter first appeared as part 1 of "Methodological Preliminaries," in *Aspects of the Theory of Syntax* (Boston: MIT Press, 1965), 3–9.

of modern general linguistics, and no cogent reason for modifying it has been offered. To study actual linguistic performance, we must consider the interaction of a variety of factors, of which the underlying competence of the speaker-hearer is only one. In this respect, study of language is no different from empirical investigation of other complex phenomena.

We thus make a fundamental distinction between *competence* (the speaker-hearer's knowledge of his language) and *performance* (the actual use of language in concrete situations). Only under the idealization set forth in the preceding paragraph is performance a direct reflection of competence. In actual fact, it obviously could not directly reflect competence. A record of natural speech will show numerous false starts, deviations from rules, changes of plan in mid-course, and so on. The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance. Hence, in the technical sense, linguistic theory is mentalistic, since it is concerned with discovering a mental reality underlying actual behavior.<sup>1</sup> Observed use of language or hypothesized dispositions to respond, habits, and so on, may provide evidence as to the nature of this mental reality, but surely cannot constitute the actual subject matter of linguistics, if this is to be a serious discipline. The distinction I am noting here is related to the *langue-parole* distinction of Saussure; but it is necessary to reject his concept of *langue* as merely a systematic inventory of items and to return rather to the Humboldtian conception of underlying competence as a system of generative processes.<sup>2</sup>

A grammar of a language purports to be a description of the ideal speaker-hearer's intrinsic competence. If the grammar is, furthermore, perfectly explicit—in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution—we may (somewhat redundantly) call it a *generative grammar*.

A fully adequate grammar must assign to each of an infinite range of sentences a structural description indicating how this sentence is understood by the ideal speaker-hearer. This is the traditional problem of descriptive linguistics, and traditional grammars give a wealth of information concerning structural descriptions of sentences. However, valuable as they obviously are, traditional grammars are deficient in that they leave unexpressed many of the basic regularities of the language with which they are concerned. This fact is particularly clear on the level of syntax, where no traditional or structuralist grammar goes beyond classification of particular examples to the stage of formulation of generative rules on any significant scale. An analysis of the best existing grammars will quickly reveal that this is a defect of principle, not just a matter of empirical detail or logical preciseness.

Nevertheless, it seems obvious that the attempt to explore this largely uncharted territory can most profitably begin with a study of the kind of structural information presented by traditional grammars and the kind of linguistic processes that have been exhibited, however informally, in these grammars.<sup>3</sup>

The limitations of traditional and structuralist grammars should be clearly appreciated. Although such grammars may contain full and explicit lists of exceptions and irregularities, they provide only examples and hints concerning the regular and productive syntactic processes. Traditional linguistic theory was not unaware of this fact. For example, James Beattie remarks that

Languages, therefore, resemble men in this respect, that, though each has peculiarities, whereby it is distinguished from every other, yet all have certain qualities in common. The peculiarities of individual tongues are explained in their respective grammars and dictionaries. Those things, that all languages have in common, or that are necessary to every language, are treated of in a science, which some have called *Universal or Philosophical* grammar.<sup>4</sup>

Somewhat earlier, Du Marsais defines universal and particular grammar in the following way:

Il y a dans la grammaire des observations qui conviennent à toutes les langues; ces observations forment ce qu'on appelle la grammaire générale: telles sont les remarques que l'on a faites sur les sons articulés, sur les lettres qui sont les signes de ces sons; sur la nature des mots, et sur les différentes manières dont ils doivent être ou arrangés ou terminés pour faire un sens. Outre ces observations générales, il y en a qui ne sont propres qu'à une langue particulière; et c'est ce qui forme les grammaires particulières de chaque langue.<sup>5</sup>

Within traditional linguistic theory, furthermore, it was clearly understood that one of the qualities that all languages have in common is their "creative" aspect. Thus an essential property of language is that it provides the means for expressing indefinitely many thoughts and for reacting appropriately in an indefinite range of new situations.<sup>6</sup> The grammar of a particular language, then, is to be supplemented by a universal grammar that accommodates the creative aspect of language use and expresses the deep-seated regularities which, being universal, are omitted from the grammar itself. Therefore it is quite proper for a grammar to discuss only exceptions and irregularities in any detail. It is only when supplemented by a universal grammar that the grammar of a language provides a full account of the speaker-hearer's competence.

Modern linguistics, however, has not explicitly recognized the necessity for supplementing a "particular grammar" of a language by a universal gram-

mar if it is to achieve descriptive adequacy. It has, in fact, characteristically rejected the study of universal grammar as misguided; and, as noted before, it has not attempted to deal with the creative aspect of language use. It thus suggests no way to overcome the fundamental descriptive inadequacy of structuralist grammars.

Another reason for the failure of traditional grammars, particular or universal, to attempt a precise statement of regular processes of sentence formation and sentence interpretation lay in the widely held belief that there is a “natural order of thoughts” that is mirrored by the order of words. Hence, the rules of sentence formation do not really belong to grammar but to some other subject in which the “order of thoughts” is studied. Thus in the *Grammaire générale et raisonnée* it is asserted that, aside from figurative speech, the sequence of words follows an “ordre naturel,” which conforms “à l’expression naturelle de nos pensées.”<sup>7</sup> Consequently, few grammatical rules need be formulated beyond the rules of ellipsis, inversion, and so on, which determine the figurative use of language. The same view appears in many forms and variants. To mention just one additional example, in an interesting essay devoted largely to the question of how the simultaneous and sequential array of ideas is reflected in the order of words, Diderot concludes that French is unique among languages in the degree to which the order of words corresponds to the natural order of thoughts and ideas.<sup>8</sup> Thus “quel que soit l’ordre des termes dans une langue ancienne ou moderne, l’esprit de l’écrivain a suivi l’ordre didactique de la syntaxe française”<sup>9</sup>; “Nous disons les choses en français, comme l’esprit est forcé de les considérer en quelque langue qu’on écrive.”<sup>10</sup> With admirable consistency he goes on to conclude that “notre langue *pédestre* a sur les autres l’avantage de l’utile sur l’agréable”<sup>11</sup>; thus French is appropriate for the sciences, whereas Greek, Latin, Italian, and English “sont plus avantageuses pour les lettres.” Moreover,

le bons sens choisirait la langue française; mais . . . l’imagination et les passions donneront la préférence aux langues anciennes et à celles de nos voisins . . . il faut parler français dans la société et dans les écoles de philosophie; et grec, latin, anglais, dans les chaires et sur les théâtres; . . . notre langue sera celle de la vérité, si jamais elle revient sur la terre; et . . . la grecque, la latine et les autres seront les langues de la fable et du mensonge. Le français est fait pour instruire, éclairer et convaincre; le grec, le latin, l’italien, l’anglais, pour persuader, émouvoir et tromper: parlez grec, latin, italien au peuple; mais parlez français au sage.<sup>12</sup>

In any event, insofar as the order of words is determined by factors independent of language, it is not necessary to describe it in a particular or universal grammar, and we therefore have principled grounds for excluding an

explicit formulation of syntactic processes from grammar. It is worth noting that this naïve view of language structure persists to modern times in various forms, for example, in Saussure's image of a sequence of expressions corresponding to an amorphous sequence of concepts or in the common characterization of language use as merely a matter of use of words and phrases.<sup>13</sup>

But the fundamental reason for this inadequacy of traditional grammars is a more technical one. Although it was well understood that linguistic processes are in some sense "creative," the technical devices for expressing a system of recursive processes were simply not available until much more recently. In fact, a real understanding of how a language can (in Humboldt's words) "make infinite use of finite means" has developed only within the last thirty years, in the course of studies in the foundations of mathematics. Now that these insights are readily available it is possible to return to the problems that were raised, but not solved, in traditional linguistic theory, and to attempt an explicit formulation of the "creative" processes of language. There is, in short, no longer a technical barrier to the full-scale study of generative grammars.

Returning to the main theme, by a generative grammar I mean simply a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences. Obviously, every speaker of a language has mastered and internalized a generative grammar that expresses his knowledge of his language. This is not to say that he is aware of the rules of the grammar or even that he can become aware of them, or that his statements about his intuitive knowledge of the language are necessarily accurate. Any interesting generative grammar will be dealing, for the most part, with mental processes that are far beyond the level of actual or even potential consciousness; furthermore, it is quite apparent that a speaker's reports and viewpoints about his behavior and his competence may be in error. Thus a generative grammar attempts to specify what the speaker actually knows, not what he may report about his knowledge. Similarly, a theory of visual perception would attempt to account for what a person actually sees and the mechanisms that determine this rather than his statements about what he sees and why, though these statements may provide useful, in fact, compelling evidence for such a theory.

To avoid what has been a continuing misunderstanding, it is perhaps worthwhile to reiterate that a generative grammar is not a model for a speaker or a hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker-hearer. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we

say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation. These questions belong to the theory of language use—the theory of performance. No doubt, a reasonable model of language use will incorporate, as a basic component, the generative grammar that expresses the speaker-hearer’s knowledge of the language; but this generative grammar does not, in itself, prescribe the character or functioning of a perceptual model or a model of speech production.<sup>14</sup>

Confusion over this matter has been sufficiently persistent to suggest that a terminological change might be in order. Nevertheless, I think that the term “generative grammar” is completely appropriate, and have therefore continued to use it. The term “generate” is familiar in the sense intended here in logic, particularly in Post’s theory of combinatorial systems. Furthermore, “generate” seems to be the most appropriate translation for Humboldt’s term *erzeugen*, which he frequently uses, it seems, in essentially the sense here intended. Since this use of the term “generate” is well established both in logic and in the tradition of linguistic theory, I can see no reason for a revision of terminology.