Chapter one

SETTING

1. Grammar

The goal of the descriptive study of a language is the construction of a grammar. We may think of a language as a set of sentences, each with an ideal phonetic form and an associated intrinsic semantic interpretation. The grammar of the language is the system of rules that specifies this sound-meaning correspondence.

The speaker produces a signal with a certain intended meaning; the hearer receives a signal and attempts to determine what was said and what was intended. The performance of the speaker or hearer is a complex matter that involves many factors. One fundamental factor involved in the speaker-hearer's performance is his knowledge of the grammar that determines an intrinsic connection of sound and meaning for each sentence. We refer to this knowledge—for the most part, obviously, unconscious knowledge—as the speaker-hearer's "competence." Competence, in this sense, is not to be confused with performance. Performance, that is, what the speaker-hearer actually does, is based not only on his knowledge of the language, but on many other factors as well—factors such as memory restrictions, inattention, distraction, nonlinguistic knowledge and beliefs, and so on. We may, if we like, think of the study of competence as the study of the potential performance of an idealized speaker-hearer who is unaffected by such grammatically irrelevant factors.

We use the term "grammar" with a systematic ambiguity. On the one hand, the term refers to the explicit theory constructed by the linguist and proposed as a description of the speaker's competence. On the other hand, we use the term to refer to this competence itself. The former usage is familiar; the latter, though perhaps less familiar, is equally appropriate. The person who has acquired knowledge of a language has internalized a system of rules that determines sound-meaning connections for indefinitely many sentences. Of course, the person who knows a language perfectly has little or no conscious knowledge of the rules that he uses constantly in speaking or hearing, writing or reading, or internal monologue. It is this system of rules that enables him to produce and interpret sentences that he has never before encountered. It is an important fact, too often overlooked, that in normal, everyday discourse one understands and produces new utterances with no awareness of novelty or innovation, although these normal utterances are similar to those previously produced or encountered only in that they are formed and interpreted by the same grammar, the same internalized system of rules. It is important to emphasize that

there is no significant sense of "generalization" in which these new utterances can be described as generalizations from earlier experience, and no sense of the term "habit" in which the normal use of language can be described as some kind of "habit system" or as "habitual behavior." We cannot, in other words, characterize the internalized, mentally represented system of rules that we call the "grammar" in terms of any other significant concept of psychology.

To summarize, then, we use the term "grammar" to refer both to the system of rules represented in the mind of the speaker-hearer, a system which is normally acquired in early childhood and used in the production and interpretation of utterances, and to the theory that the linguist constructs as a hypothesis concerning the actual internalized grammar of the speaker-hearer. No confusion should result from this standard usage if the distinction is kept in mind.

2. *Linguistic universals*

General linguistics attempts to develop a theory of natural language as such, a system of hypotheses concerning the essential properties of any human language. These properties determine the class of possible natural languages and the class of potential grammars for some human language. The essential properties of natural language are often referred to as "linguistic universals." Certain apparent linguistic universals may be the result merely of historical accident. For example, if only inhabitants of Tasmania survive a future war, it might be a property of all then existing languages that pitch is not used to differentiate lexical items. Accidental universals of this sort are of no importance for general linguistics, which attempts rather to characterize the range of possible human languages. The significant linguistic universals are those that must be assumed to be available to the child learning a language as an a priori, innate endowment. That there must be a rich system of a priori properties—of essential linguistic universals—is fairly obvious from the following empirical observations. Every normal child acquires an extremely intricate and abstract grammar, the properties of which are much underdetermined by the available data. This takes place with great speed, under conditions that are far from ideal, and there is little significant variation among children who may differ greatly in intelligence and experience. The search for essential linguistic universals is, in effect, the study of the a priori *faculté de langage* that makes language acquisition possible under the given conditions of time and access to data.

It is useful to divide linguistic universals roughly into two categories. There are, first of all, certain "formal universals" that determine the structure of grammars and the form and organization of rules. In addition, there are "substantive universals" that define the sets of elements that may figure in particular grammars. For example, the theory of transformational generative grammar proposes certain formal universals regarding the kinds of rules that can appear in a grammar, the kinds of structures on which they may operate, and the ordering conditions under which these rules may apply. We shall study these questions in detail, in connection with the phonological component of a generative grammar. Similarly, general linguistic theory might propose, as substantive universals, that the lexical items of any language are assigned to fixed categories such as noun, verb, and adjective, and that phonetic transcriptions must make use of a particular, fixed set of phonetic features. The latter topic, once again, will occupy us in this book. We will be concerned with the theory of "universal phonetics," that part of general linguistics that specifies the class of "possible phonetic representations" of sentences by determining the universal set of phonetic features and the conditions on their possible combinations. The phonetic form of each sentence in each language is drawn from this class of possible phonetic representations.

3. *Phonetic representations*

What exactly is a phonetic representation? Suppose that universal phonetics establishes that utterances are sequences of discrete segments, that segments are complexes of a particular set of phonetic features, and that the simultaneous and sequential combinations of these features are subject to a set of specific constraints. For example, universal phonetics may provide us with the feature "consonantal," which distinguishes [+consonantal] phonetic segments such as [p], [t], [θ], [s], [š] from [−consonantal] phonetic segments such as [u], [i], [a]; and the feature "strident," which distinguishes [+strident] segments such as [s] and [š] from [−strident] segments such as [p], [t], and [θ]. Among the "simultaneous constraints" of universal phonetics would be the condition that no phonetic segment can be both [−consonantal] and [+strident]; the feature "strident" does not provide a further classification of the category of [−consonantal] segments. Among the "sequential constraints" might be certain conditions that assign a maximal length to a sequence of [+consonantal] phonetic segments, that is, to a consonant cluster. There will be many other constraints of both sorts, and they must be met by each phonetic representation in each language.

More specifically, a phonetic representation has the form of a two-dimensional matrix in which the rows stand for particular phonetic features; the columns stand for the consecutive segments of the utterance generated; and the entries in the matrix determine the status of each segment with respect to the features. In a full phonetic representation, an entry might represent the degree of intensity with which a given feature is present in a particular segment; thus, instead of simply subdividing segments into [+strident] and [−strident], as in the example just given, the entries in the row corresponding to the feature "strident" might indicate degrees along a differentiated scale of "stridency." The phonetic symbols [p], [t], [θ], [i], [u], etc., are simply informal abbreviations for certain feature complexes; each such symbol, then, stands for a column of a matrix of the sort just described.

To recapitulate, the phonetic representation of an utterance in a given language is a matrix with rows labeled by features of universal phonetics. The grammar of the language assigns to this phonetic representation a "structural description" that indicates how it is to be interpreted, ideally, in this language. More generally, we may say that the grammar of each language assigns a structural description to each member of the universal class of possible phonetic representations. For example, the grammar of every language will assign structural descriptions to phonetic representations such as (1) and (2):[1]

(1) ilvyɛ̃dradəmɛ̃ ("il viendra demain")

(2) hiylkʌm+təmarə ("he'll come tomorrow")

[1] We omit much phonetic detail that should be specified in universal representations but that is irrelevant to the exposition here. This is the course we will generally follow in discussing particular examples. In the representation (2), and in other representations in this chapter, we include the "boundary symbol" +, which can be taken as specifying a certain type of transition between phonetic elements. Actually, however, we will suggest later that boundary symbols do not appear in phonetic representations.

The grammar of English will assign to (1) a structural description indicating that it is not a sentence of English at all, and to (2) a structural description that specifies the elements of which it is composed on the various linguistic levels, the manner of their organization, the interrelations of these abstract representations, and so on. The grammar of French will supply this information for (1), and will designate (2) as a nonsentence. Many elements of the class of possible phonetic representations will be designated as "semi-grammatical sentences," not well-formed but nevertheless interpretable by analogy to well-formed sentences in ways that are, for the moment, not well understood.[2]

4. Components of a grammar

The class of possible phonetic representations is of course infinite. Similarly, the class of phonetic representations designated as well-formed sentences in each human language is infinite. No human language has a limit on the number of sentences that are properly formed and that receive a semantic interpretation in accordance with the rules of this language. However, the grammar of each language must obviously be a finite object, realized physically in a finite human brain. Therefore, one component of the grammar must have a recursive property; it must contain certain rules that can be applied indefinitely often, in new arrangements and combinations, in the generation (specification) of structural descriptions of sentences. Every language, in particular, contains processes that permit a sentence to be embedded within another sentence, as the English sentence *John left* is embedded in the sentence *I was surprised that John left*. These processes can apply indefinitely often to form sentences of arbitrary complexity. For example, the sentence *I was surprised that John left* can itself be embedded in the context *Bill expected* ——, giving, finally, *Bill expected me to be surprised that John left*, after various obligatory modifications have taken place. There is no limit to the number of applications of such processes; with each further application, we derive a well-formed sentence with a definite phonetic and semantic interpretation.

The part of a grammar which has this recursive property is the "syntactic component," the exact form of which will not concern us here.[3] We will, however, make certain assumptions about the abstract objects generated by the syntactic component, that is, about the "syntactic descriptions" that can be formed by the application of its rules.

The syntactic component of a grammar assigns to each sentence a "surface structure" that fully determines the phonetic form of the sentence. It also assigns a far more abstract "deep structure" which underlies and partially determines the surface structure but is otherwise irrelevant to phonetic interpretation, though it is of fundamental significance for semantic interpretation. It is important to bear in mind that deep structures are very different from the surface structures to which we will restrict our attention and that they provide a great deal of information not represented in surface structures.

To recapitulate, a grammar contains a syntactic component which is a finite system of rules generating an infinite number of syntactic descriptions of sentences. Each such syntactic description contains a deep structure and a surface structure that is partially determined by the deep structure that underlies it. The semantic component of the grammar

[2] For discussion of this matter, which we will exclude from consideration henceforth, see Section IV of Fodor and Katz (1964), and pages 148 ff. of Chomsky (1965), as well as many other references.

[3] For recent discussion, see Katz and Postal (1964) and Chomsky (1965).

is a system of rules that assigns a semantic interpretation to each syntactic description, making essential reference to the deep structure and possibly taking into account certain aspects of surface structure as well. The phonological component of the grammar assigns a phonetic interpretation to the syntactic description, making reference only to properties of the surface structure, so far as we know. The structural description assigned to a sentence by the grammar consists of its full syntactic description, as well as the associated semantic and phonetic representations. Thus the grammar generates an infinite number of sentences, each of which has a phonetic and semantic representation; it defines an infinite sound-meaning correspondence, this correspondence being mediated by the abstract syntactic component and the structures it generates.

We are not concerned here with deep structures and the rules that generate them, the rules that relate them to surface structures, or the rules that assign semantic interpretations to syntactic descriptions. We are limiting our attention to surface structures, phonetic representations, and the rules that assign a phonetic representation (possibly several phonetic representations, in the case of free variation) to each surface structure.

5. Surface structures

The surface structures generated by the syntactic component have the following characteristics. Each consists of a string of minimal elements that we will call "formatives." Each formative is assigned to various categories that determine its abstract underlying form, the syntactic functions it can fulfill, and its semantic properties. For example, the formative *boy* will belong to the category of elements with initial voiced stops,[4] to the category "noun," to the category "animate," to the category "male," etc. This information about formatives will be presented in a "lexicon," which forms part of the syntactic component of the grammar. The organization of the lexicon will not concern us here; we simply assume that the full categorization of each formative is represented in the surface structure. In fact, we may think of the lexical entry of a formative as nothing other than a list of the categories to which it belongs. The categories are sometimes called "features." We will refer, as we proceed, to phonological, syntactic, and semantic features.

The surface structure must indicate how the string of formatives it contains is subdivided into "phrases," each phrase being a certain continuous substring of the string of formatives. The analysis of strings into phrases is a "proper bracketing" in the sense that phrases can overlap only if one is contained in the other. Thus, if A, B, C are formatives, the surface structure of the string ABC cannot specify AB as a phrase and BC as a phrase, for the string may be bracketed either as $((AB)C)$ or as $(A(BC))$ but not in both ways simultaneously.

The phrases furthermore are assigned to certain categories, and this information may be represented by putting labels on the brackets. Take, for example, the sentence (3):

(3) *we established telegraphic communication*

In (3), the string underlying *we* is assigned to the same category as the string underlying

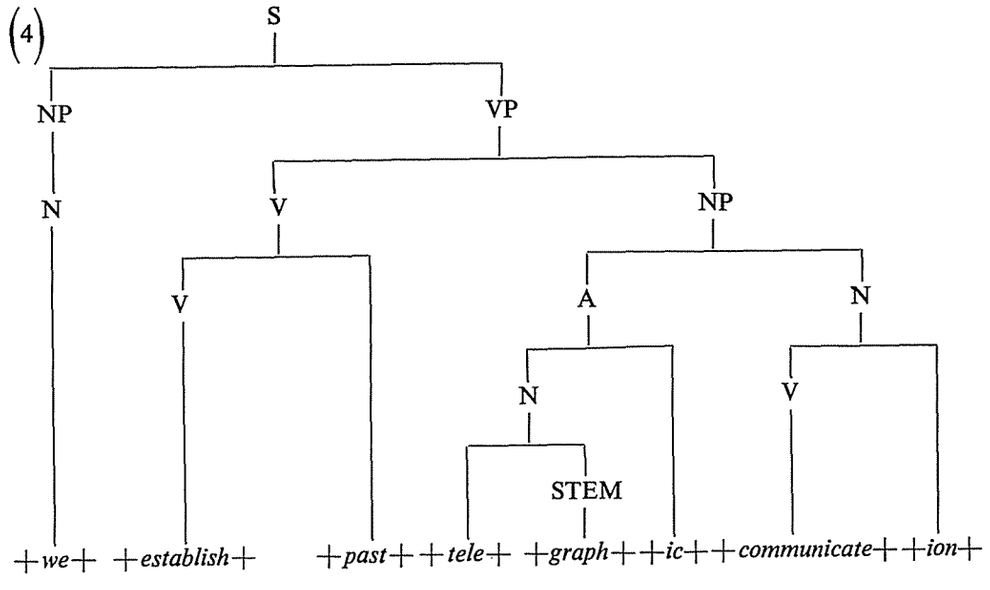
[4] This underlying representation will be abstract in a sense that we will later describe in detail. For example, although the formative *boy* is always represented phonetically with a back vowel, we will present evidence showing that it should be represented in surface structure—that is, before the phonological rules apply—with a front vowel.

telegraphic communication, namely, to the category "noun phrase." Similarly, the other phrases are assigned to certain universal categories.

We will make the empirical assumption that the surface structure of a sentence is precisely a proper bracketing of a string of formatives, with the bracketed substrings (the phrases) assigned to categories selected from a certain fixed universal set of categories. The complete string is assigned to the category "sentence"(S); the other phrases are also assigned to categories that are provided by general linguistic theory, such as the categories "noun phrase" (NP) and "verb phrase" (VP). These universal categories are on a par with the phonetic categories (bilabial closure, frontness, etc.) provided by universal phonetic theory. As we noted earlier, the categories of universal phonetic theory determine a certain infinite class of possible phonetic representations from which the phonetic forms of sentences of any human language are drawn. Similarly, the universal set of phrase categories (NP, VP, etc.), together with the universal lexical categories (noun, verb, adjective) and the universal lexical features that define the class of "possible formatives," provides us with an infinite class of possible surface structures, from which the surface structures of sentences of any particular language are drawn. In other words, general linguistics should provide definitions, in terms independent of any particular language, for the notions "possible phonetic representation" and "possible surface structure." The grammar of each language relates phonetic representations to surface structures in a specific way; and, furthermore, it relates surface structures to deep structures, and, indirectly, to semantic interpretations, in ways that are beyond the scope of our present study.

To give a concrete example, the grammar of English might assign to the sentence (3) a surface structure which can be represented in the equivalent forms (4) and (5):[5]

(4)



(5) $[_{S}\,[_{NP}\,[_{N}+we+]_{N}\,]_{NP}\,[_{VP}\,[_{V}\,[_{V}+establish+]_{V}+past+]_{V}\,[_{NP}\,[_{A}\,[_{N}+tele+\ [_{STEM}+graph+]_{STEM}\,]_{N}+ic+]_{A}\,[_{N}\,[_{V}+communicate+]_{V}+ion+]_{N}\,]_{NP}\,]_{VP}\,]_{S}$

[5] Once again (see note 1), we omit details which are irrelevant here. We assume, for the purposes of this example, that the formatives are *we*, *establish*, *past*, *tele*, *graph*, *ic*, *communicate*, *ion*. The node labeled A represents the lexical category "adjective"; the other labels have been mentioned previously.

The interpretation of the notational devices used in (4) and (5) should be obvious. We intend these representations to indicate that the formative *we* is both an N and an NP, the formative *establish* a V, the formative string *tele graph* an N, the formative string *tele graph ic communicate ion* an NP, the full string an S, etc.[6] Furthermore, each formative has an analysis as a set of intersecting categories, in a way that we shall specify in more detail below. The + symbols represent formative boundaries which, by convention, automatically mark the beginning and end of each formative.

5.1. *LEXICAL AND PHONOLOGICAL REPRESENTATIONS*

To recapitulate, we presuppose, for our description of English sound patterns, a grammar with a syntactic component that assigns to each sentence a surface structure such as (4)-(5), that is, a proper labeled bracketing of a string of formatives. Our main concern here will be the "phonological component," that is, the system of rules that applies to a surface structure and assigns to it a certain phonetic representation drawn from the universal class provided by general linguistic theory. In particular, the phonological rules of English must assign to the surface structure (4)-(5) a phonetic representation much like (6):

(6) wīyəstæ̂blišt+tèləgræ̂fik+kəmyůwnəkéyšən

The phonetic representation (6), corresponding to the underlying surface structure (4)-(5), is a feature matrix of the sort described earlier. In the surface structure, the individual formatives (for example, the lexical formatives *we*, *establish*, *tele*, *graph*, *communicate*, and the grammatical formatives *past*, *ic*, *ion*) will themselves be represented as feature matrices of an abstract sort, and we must now say a few words about this kind of representation. We shall distinguish between "lexical representations" and "phonological representations." We shall use the term "lexical representation" in reference to formatives which are provided directly by the lexicon, i.e., the lexical formatives as well as certain grammatical formatives which happen to appear in lexical entries. There may be other grammatical formatives introduced directly by the syntactic rules themselves. Thus the syntactic rules and the lexicon, applied in a manner that does not concern us here, provide for each utterance a representation as a string of formatives with surface structure.

Notice, however, that the surface structure must meet two independent conditions: first, it must be appropriate for the rules of phonological interpretation; second, it must be "syntactically motivated," that is, it must result from the application of independently motivated syntactic rules. Thus we have two concepts of surface structure: input to the phonological component and output of the syntactic component. It is an empirical question whether these two concepts coincide. In fact, they do coincide to a very significant degree, but there are also certain discrepancies. These discrepancies, some of which we discuss as we proceed, indicate that the grammar must contain certain rules converting the surface structures generated by the syntactic component into a form appropriate for use by the phonological component. In particular, if a linguistic expression reaches a certain level of complexity, it will be divided into successive parts that we will call "phonological phrases," each of which is a maximal domain for phonological processes. In simple cases the whole sentence is a single phonological phrase; in more complex cases the sentence may be reanalyzed as a sequence of phonological phrases. The analysis into phonological phrases

[6] Since in representations such as (4) the category labels are placed above the elements in the string that belong to these categories, one frequently speaks of the category as "dominating" a string or a part of a string. Thus, with respect to (4), we will say both that *we* "is an" N and that *we* "is dominated by" N.

depends in part on syntactic structure, but it is not always syntactically motivated in the sense just mentioned. If the syntactic component were to be connected to an orthographic rather than a phonetic output system, the reanalysis into phonological phrases would be unnecessary. Writers, unlike speakers, do not run out of breath, and are not subject to other physiological constraints on output that require an analysis into phonological phrases.

In addition to a reanalysis into phonological phrases in complex cases, the "readjustment rules" relating syntax to phonology make various other modifications in surface structures. It seems that in general these modifications involve elimination of structure, that is, deletion of nodes in representations such as (4) or of paired brackets in representations such as (5). One can easily imagine why this should be so. Reasoning along lines suggested in Miller and Chomsky (1963, Part 2), let us suppose that perception involves a two-stage memory. The first stage is a short-term system quite limited in capacity and operating in real time in the sense that it must remain available for receiving the incoming signal, and the second stage is a very large system that operates on information supplied to it by the short-term real-time system. The short-term first stage must provide an initial analysis of the signal that is just sufficient in detail to permit the second-stage system to derive the deep structure and semantic interpretation. We might expect a language to be so designed that a very superficial analysis into phrases can be performed by a system with limited memory and heavy restrictions on access. To relate this speculation to the discussion of surface structure, it appears that the syntactic component of the grammar generates a surface structure Σ which is converted, by readjustment rules that mark phonological phrases and delete structure, to a still more superficial structure Σ'. The latter then enters the phonological component of the grammar. We might speculate, then, that a first stage of perceptual processing involves the recovery of Σ' from the signal using only the restricted short-term memory, and that a second stage provides the analysis into Σ and the deep structure that underlies it. From this point of view, it would be natural to suppose that the readjustment rules that form Σ' from Σ will have the effect of reducing structure. It is, incidentally, worthy of note that the transformations that form surface structures from deep structures also characteristically have the effect of reducing structure, in a sense which can be made precise.[7]

Let us return now to our discussion of lexical and phonological representations. We have used the term "lexical representation" to refer to the representation of formatives provided by the lexicon. As we have stated, however, the structures generated through the interaction of syntactic and lexical rules are not quite appropriate, in certain cases, for the application of the rules of the phonological component. They must be modified by certain readjustment rules (of a sort to which we will return in Chapter Eight, Section 6.5, noting, however, that our investigation of the effects of surface structure on phonetic representation has not yet reached a level of depth and complexity that requires a detailed, formal analysis of these processes).[8] These readjustment rules may somewhat modify the labeled bracketing of surface structure. They may also construct new feature matrices for certain strings of lexical and grammatical formatives. To take an obvious example, the verb *sing* will appear in the lexicon as a certain feature matrix, as will the verb *mend*. Using letters of the alphabet as informal abbreviations for certain complexes of features, i.e., certain columns of a feature matrix, we can represent the syntactically generated surface structure underlying the

[7] See Miller and Chomsky (1963). See also Ross (1967) for further relevant observations of a different sort on reduction of structure under transformations.

[8] See Bierwisch (1966) for a very interesting study of readjustment rules of the sort mentioned here.

forms *sang* and *mended* as $[_V\ [_V sing]_V\ past]_V$ and $[_V\ [_V mend]_V\ past]_V$, respectively, where *past* is a formative with an abstract feature structure introduced by syntactic rules. The readjustment rules would replace *past* by *d*, as a general rule; but, in the case of *sang*, would delete the item *past* with the associated labeled brackets, and would add to the *i* of *sing* a feature specification indicating that it is subject to a later phonological rule which, among other things, happens to convert *i* to æ. Designating this new column as *, the readjustment rules would therefore give the forms $[_V s*ng]_V$ and $[_V\ [_V mend]_V\ d]_V$, respectively. We shall refer to this representation—and in general to the representation given by the application of all readjustment rules—as the "phonological representation."

Other terms that might have been used in place of the terms just proposed are "morphophonemic representation" or "systematic phonemic representation." We have avoided these terms, however, because of the technical meaning they have been given in various theories of sound structure developed in modern linguistics. The term "morphophonemic representation" seems to us appropriate only if there is another linguistically significant level of representation, intermediate in "abstractness" between lexical (phonological) and phonetic and meeting the conditions placed on "phonemic representation" in modern structural linguistics. We feel, however, that the existence of such a level has not been demonstrated and that there are strong reasons to doubt its existence.[9] We will make no further mention of "phonemic analysis" or "phonemes" in this study and will also avoid terms such as "morphophonemic" which imply the existence of a phonemic level. Notice that the issue in this case is not terminological but rather substantive; the issue is whether the rules of a grammar must be so constrained as to provide, at a certain stage of generation, a system of representation meeting various proposed conditions. The references in note 9 explain our position, and we will say no more about the matter here.

5.2. *ON THE ABSTRACTNESS OF LEXICAL REPRESENTATIONS*

We have said that the underlying representations, lexical as well as phonological, are abstract as compared with phonetic representations, although both are given in terms of phonetic features. The meaning of this remark will become clearer as we proceed. There is, however, one very obvious sense in which the underlying representations are more abstract than the phonetic representations. Consider, for example, the word *telegraph*. This has several different variants in actual phonetic representations:[10]

(7) tel^{1}əgræf^{3}[11] (in isolation)

(8) tel^{3}əgr^{1}æf (in the context —— *ic*; i.e., *telegraphic*)

(9) tel^{1}əgrəf (in the context ——*y*; i.e., *telegraphy*)

It is quite obvious, however, that this phonetic variation is not fortuitous—it is not of the

[9] We have presented our reasons for doubting the existence of a phonemic level, in the sense of modern linguistics, in various places. See Halle (1959), Chomsky (1964, 1966b), and Chomsky and Halle (1965), as well as Postal (1962, 1968), for arguments that seem to us fully convincing.

[10] Notice that in the sentence (6) it has still another representation because of the stress modifications that take place in that context.

[11] Stress levels are indicated here and throughout by numerals, with "1" representing primary stress, "2" representing secondary stress, etc. (See also note 3 in Chapter Two on this subject.)