THE SOUND PATTERN OF ENGLISH

NOAM CHOMSKY
MORRIS HALLE

The MIT Press
Cambridge, Massachusetts
London, England

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

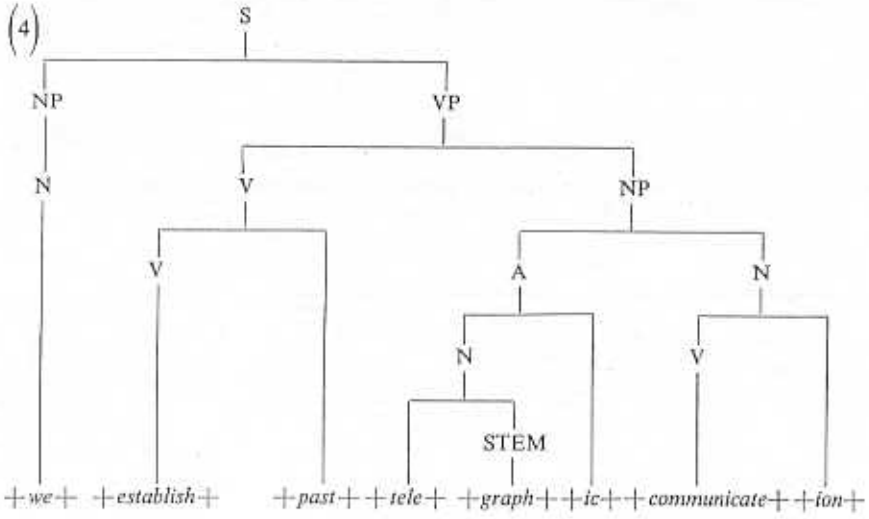
[4] This underlying representation will be abstract in a sense that we will later describe in detail. For example, although the formative *boy* is always represented phonetically with a back vowel, we will present evidence showing that it should be represented in surface structure—that is, before the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

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) tɛlə$\overset{1}{\text{græ}}\overset{3}{\text{f}}$[11] (in isolation)

(8) tɛlə$\overset{3}{\text{gr}}\overset{1}{\text{æf}}$ (in the context —— *ic*; i.e., *telegraphic*)

(9) tə$\overset{1}{\text{l}}$ɛgrəf (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.)

same type as the variation between *I* and *we*, which depends on specific assignment of the latter to the category of plurality. Given the grammar of English, if we delete specific reference to the item *we*, there is no way to predict the phonetic form of the plural variant of *I*. On the other hand, the rules for English grammar certainly do suffice to determine the phonetic variation of *telegraph* without specific mention of this lexical item, just as they suffice to predict the regular variation between *cat* and *cats* without specifically mentioning the plural form. It is quite obvious that English grammar is complicated by the fortuitous variation between *I* and *we* but not by the totally predictable variation between *cat* and *cats*. Similarly, the grammar would be more complicated if *telegraph* did *not* undergo precisely the variation in (7)–(9): if, for example, it had one phonetic form in all contexts, or if it had the form (7) in the context ——*ic*, (8) in the context ——*y*, and (9) in isolation.

In short, the phonetic variation of *telegraph* in certain contexts is not an idiosyncratic property of this particular lexical item but is rather a matter of general rule, applying to many other lexical items as well. Regular variations such as this are not matters for the lexicon, which should contain only idiosyncratic properties of items, properties not predictable by general rule. The lexical entry for *telegraph* must contain just enough information for the rules of English phonology to determine its phonetic form in each context; since the variation is fully determined, the lexical entry must contain no indication of the effect of context on the phonetic form. In fact, as we shall see, the lexical representation for the word *telegraph* should be (10), where each of the symbols *t*, *e*, ... is to be understood as an informal abbreviation for a certain set of phonological categories (distinctive features):[12]

(10) +tele+græf+

Thus the lexical representation is abstract in a very clear sense; it relates to the signal only indirectly, through the medium of the rules of phonological interpretation that apply to it as determined by its intrinsic abstract representation and the surface structures in which it appears.

An analogous argument can readily be constructed for the abstract nature of the phonological representations, i.e., those representations that are determined from lexical representations by application of certain readjustment rules (and which, for the most part, are in fact identical with lexical representations).

5.3. *ANALYSIS INTO WORDS*

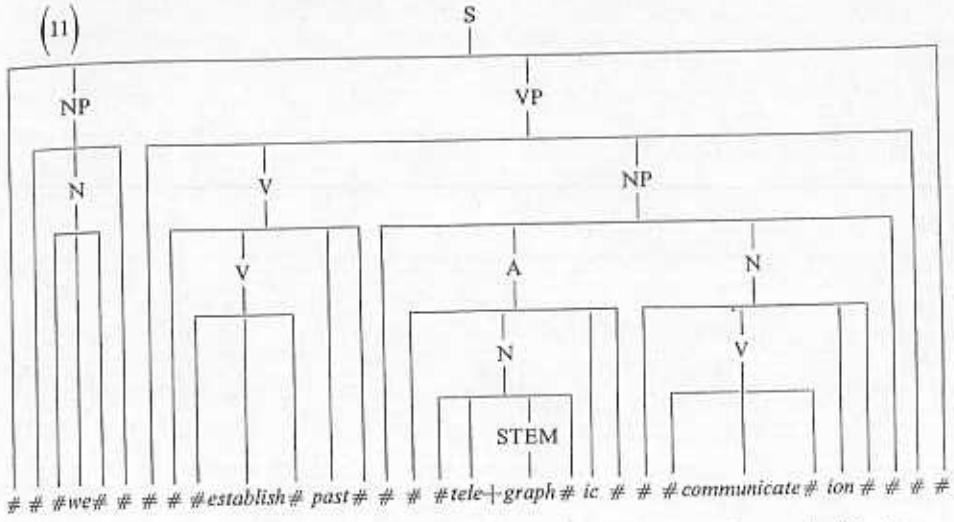
One additional aspect of surface structure is crucial for our discussion. We will see that the phonological rules fall into two very different classes. Certain of these rules apply freely to phrases of any size, up to the level of the phonological phrase; others apply only to words. We must therefore assume that the surface structure of an utterance provides an analysis into a sequence of words. For example, the sentence (3), *we established telegraphic communication*, will be analyzed by its surface structure into the four successive words *we*, *establish+past*, *tele+graph+ic*, *communicate+ion*. The rules that form surface structure (or, perhaps, the readjustment rules discussed above) must provide this information, since it is required for the correct application of the rules of the phonological component of the grammar.

As a first approximation to the problem of analysis into words, let us assume that each lexical category (e.g., noun, verb, adjective) and each category that dominates a lexical

[12] In addition, the lexical entry will provide the other idiosyncratic syntactic information represented in (4)-(5), namely, the information that *graph* is a stem and *telegraph* is a noun.

category (e.g., sentence, noun phrase, verb phrase) automatically carries a boundary symbol # to the left and to the right of the string that belongs to it (i.e., that it dominates, in tree representations such as (4), or that it brackets, in bracket representations such as (5)). Under this assumption, we replace the representation (4) by (11) and modify (5) in a corresponding way:

(11)



Let us tentatively define a word as a string of formatives (one or more) contained in the context # # —— # # and containing no occurrences of # #.[13] Thus the words in (11) are *we*, *establish#past*, *tele+graph#ic*, and *communicate#ion*, as required. The principle just stated can be regarded, tentatively, as a universal principle for the interpretation of surface structures, and, as a first approximation, it works quite well. Among the readjustment rules discussed above, there will be some that modify the representation provided by this interpretive principle in ad hoc ways. For example, we shall see that although the boundary # is appropriate in *establish#ed*, as distinct from ordinary formative boundary (which we have been representing as +), it must be replaced by ordinary formative boundary in *tele+graph#ic* and *communicate#ion*, for reasons having to do with the applicability of certain phonetic rules.

To recapitulate, the rules of syntax will generate surface structures and a universal principle of interpretation will assign the boundary symbol # in certain places. The readjustment rules will modify the surface structure in various ad hoc ways, demarcating it into phonological phrases, eliminating some structure, and replacing some occurrences of # by +. The abstract object thus constructed (which we will also refer to as a "surface structure," or, if more explicitness is necessary, a "phonological surface structure," to contrast it with the syntactic surface structure generated by the syntactic component) enters the phonological component of the grammar and is converted by the phonological rules into a phonetic representation, in ways that we will specify in detail as we proceed. Certain of the phonological rules will apply only to words; others will apply freely to strings of formatives which may be words or subparts of words, or phrases that include words.

[13] See Chapter Eight, Section 6.2, for a more careful analysis of the notion "word."

We will find it convenient to use labeled bracketing such as (5) rather than tree diagrams such as (4) and (11) for the representation of surface structure in the presentation of phonological rules. Since, by convention, every lexical category or category dominating a lexical category has # boundaries associated with it on the left and right, we will sometimes omit reference to these boundaries in the statement of rules. For example, a rule of the form (12) is to be understood as applying to the string (13):

$$(12) \qquad A \rightarrow B \ / \ X \text{——} Y]_V$$

$$(13) \qquad XAY\#]_V$$

Rule (12) states that an element of the type A is rewritten as a corresponding element of the type B when A appears in the context X —— Y (that is, with X to its left and Y to its right) and when the item in question is a verb, i.e., is dominated by V or, equivalently, is bracketed by $[_V\ ]_V$. We will make these informal specifications more precise as we proceed.

6. Summary

The phonological component is a system of rules such as (12) that relates surface structures such as (11) to phonetic representations such as (6). As we proceed in our discussion, we will propose various specific hypotheses regarding the detailed form of representations such as (11) and (6), and we will also make specific proposals concerning the system of phonological rules that assign a phonetic interpretation to each surface structure.

We have already suggested that a phonetic representation such as (6) is actually a feature matrix in which the rows correspond to a restricted set of universal phonetic categories or features (voicing, nasality, etc.) and the columns to successive segments. We will propose further that such representations are mentally constructed by the speaker and the hearer and underlie their actual performance in speaking and "understanding." We will consider the question of the relation between such phonetic representations and actual speech signals, and the steps by which such representations might be constructed by the hearer on the occasion of reception of a speech signal. We have suggested, moreover, that each formative of the surface structure can also be represented as a feature matrix interpreted in a rather similar way, with rows corresponding to the universal phonetic and grammatical categories. The formative structure is much more abstract, however; its relation to the speech signal is not as direct as that of the phonetic representation.

We will propose that the rules of the phonological component have a fixed form and a specific organization, that they apply in a fixed manner determined by the labeled bracketing of the surface structure, and that they meet various additional conditions depending on their formal relations. These we propose as universal conditions, as aspects of general linguistic theory. We will try to show how, on the basis of these assumptions, many particular phenomena of English sound structure can be explained.

With these remarks on background assumptions, we can proceed to the analysis of English sound structure and of general phonological theory.

Chapter two

A SKETCH OF ENGLISH PHONOLOGY AND PHONOLOGICAL THEORY

1. The principle of the transformational cycle and its application to English stress contours

We turn here to the problem of how a surface structure of the sort described in the preceding chapter determines a phonetic representation.

It is well known that English has complex prosodic contours involving many levels of stress and pitch[1] and intricate processes of vowel reduction. It is clear even from a superficial examination that these contours are determined in some manner by the surface structure of the utterance. Furthermore, it is natural to suppose that in general the phonetic shape of a complex unit (a phrase) will be determined by the inherent properties of its parts and the manner in which these parts are combined, and that similar rules will apply to units of different levels of complexity. These observations suggest a general principle for the application of rules of the phonological component, namely, what we shall call the principle of the "transformational cycle."[2] Regarding a surface structure as a labeled bracketing (see representation (5) in Chapter One), we assume as a general principle that the phonological rules first apply to the maximal strings that contain no brackets, and that after all relevant rules have applied, the innermost brackets are erased; the rules then reapply to maximal strings containing no brackets, and again innermost brackets are erased after this application; and so on, until the maximal domain of phonological processes is reached. In terms of the tree representation of a surface structure (see representation (4) in Chapter One), the rules apply to a string dominated by a particular node A only after they have already applied to the strings dominated by each of the nodes dominated by A.

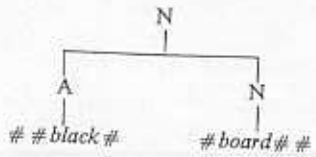
The actual operation of the transformational cycle can now be illustrated with some simple examples. It is clear, first of all, that there are at least two processes of stress

[1] As we explained in the Preface, we will have nothing to say about pitch in this study.

[2] This principle was first formulated in Chomsky, Halle, Lukoff (1956) in a slightly different but equivalent terminology. It has since been applied to phonetic study of a variety of different languages: French (Schane, 1965), Russian (Halle, 1963, Lightner, 1965a), Japanese (McCawley, 1965).

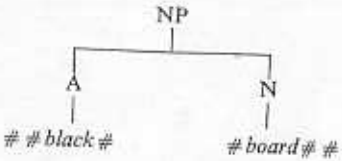
assignment in English. Thus *blackboard*,[3] with a falling stress contour, must be distinguished from *black board*, with a rising contour. The elementary constituents, *black*, an adjective, and *board*, a noun, are the same in both cases; the difference lies in the way these constituents are combined, as reflected in their different surface structures, shown here in the two notations of the preceding chapter:

(1) (a)



$[_N \# [_A \# black \#]_A [_N \# board \#]_N \#]_N$

(b)



$[_{NP} \# [_A \# black \#]_A [_N \# board \#]_N \#]_{NP}$

In case (1a), where the entire phrase belongs to the category "noun," the phonological rules must give the contour 13; in case (1b), where it belongs to the category "noun phrase," the rules must give the contour 21. According to the principle of the transformational cycle, the phonological rules apply first to the strings dominated by A and by N, the lowest-level categorial nodes of (1); in other words, the rules apply first to *black* and to *board*. In isolation, each of these would receive primary stress. We therefore might propose the rule:

(2) In monosyllables, the vowel receives primary stress.

Applying this rule to the structures of (1) and then erasing innermost brackets in accordance with the principle of the transformational cycle, we have, in the bracket notation, the representations (3a) and (3b):

(3) (a) $[_N \# \# \overset{1}{black} \# \ \# \overset{1}{board} \# \#]_N$ (b) $[_{NP} \# \# \overset{1}{black} \# \ \# \overset{1}{board} \# \#]_{NP}$

We must now apply rules that weaken the rightmost primary stress in case (3a) and that weaken the leftmost primary stress in case (3b). For many reasons, it is necessary to state the rules that determine stress contours as rules of placement of primary stress, rather than as rules of stress weakening. We will therefore formulate the rules that apply to (3) as processes that place primary stress on the leftmost and the rightmost syllables, respectively, and we will adopt the following convention: *when primary stress is placed in a certain position, then all other stresses in the string under consideration at that point are automatically weakened by one.* We can now state the following two rules:

(4) Assign primary stress to a primary-stressed vowel in the context

$$\text{——} \ldots \overset{1}{\text{V}} \ldots]_N$$

(5) Assign primary stress to a primary-stressed vowel in the context

$$\overset{1}{\text{V}} \ldots \text{——} \ldots]_{NP}$$

In rules (4) and (5), the symbol V stands for "vowel," and $\overset{1}{\text{V}}$ stands for a vowel with primary stress. The dash indicates the position of the segment to which the rule applies. Thus rule (4) assigns primary stress to a primary-stressed vowel which is *followed* by another primary-stressed vowel in a noun, and rule (5) assigns primary stress to a primary-stressed vowel which is *preceded* by another primary-stressed vowel in a noun phrase. By the convention stated above, the actual effect of these rules is to weaken the other stresses in the string to which the rule applies. Thus, applying rule (4) to (3a), we derive the representation (6a); applying rule (5) to (3b), we derive the representation (6b).

(6) (a) $\# \# \overset{1}{black} \# \# \overset{2}{board} \# \#$ (b) $\# \# \overset{2}{black} \# \# \overset{1}{board} \# \#$

We will refer to (4) as the Compound Rule and to (5) as the Nuclear Stress Rule.

It is important to observe that rules (4) and (5) make use of the bracketing given in the surface structure for their proper cyclic operation, and that the labels on the brackets, that is, the syntactic categories indicated in the surface structure, are necessary for determining the correct application of the rules.

To derive the stress contour for *blackboard*, we must apply still another rule, weakening the secondary stress on the second syllable to tertiary. This process can be formulated in the following way (with C_0 standing for a string of zero or more consonants):

(7) Assign primary stress to a primary-stressed vowel in the context

$$\text{——} \ldots \# \# C_0 \overset{2}{\text{V}} C_0 \#]_N$$

Application of rule (7) to (6a) gives the desired stress pattern 13 by the conventions established above; primary stress is placed on the first syllable, and the stress on the second syllable is automatically weakened to tertiary.

Clearly, both the Compound Rule and the Nuclear Stress Rule are of much greater generality than is indicated by the formulation we have given. Thus, rule (4) actually applies not only to compound nouns such as *blackboard*, but also to compound adjectives (*heart-broken*) and compound verbs (*air-condition*). It must therefore be extended to lexical categories in general. Similarly, the Nuclear Stress Rule applies not only to noun phrases, but to any phrase which is not a lexical category—for example, to verb phrases (*read the book*), to adjective phrases (*eager to please*), and to whole sentences (*John left*). We therefore replace rules (4) and (5) by the formulations (8) and (9):

(8) Assign primary stress to a primary-stressed vowel in the context

$$\text{——} \ldots \overset{1}{\text{V}} \ldots]_{NAV}$$

(9) Assign primary stress to a primary-stressed vowel in the context

$$\overset{1}{\text{V}} \ldots \text{——} \ldots]_\alpha$$

[3] There are various conventions in use for marking stress, which, at least in part, appear to differ in factual content. We return to this matter later. Here, as mentioned in Chapter One, note 11, in place of the conventional symbols ˈ, ˌ, ˈ, ˉ for primary, secondary, tertiary, and quaternary (zero) stress, respectively, we will simply use numerals, starting with 1 for primary stress. It should be kept in mind that the numbers go down as the stress goes up, admittedly a disadvantage of this notation. To minimize confusion, we will speak of strengthening and weakening stress, rather than of increasing and decreasing it.

where $]_a$ stands for a bracket with any label except N, A, or V. We can make the notion "except" precise in a very simple way, namely, by requiring that the rules (8) and (9) apply in the order given. We can then take α in (9) to be simply a variable ranging over all categories. If rule (8) has applied, the resulting string will contain just one primary stress and thus will not fit the required context for (9). Therefore (9) will never apply when $\alpha = N, A, or V$.

Using familiar notations, we can now formulate the Compound and Nuclear Stress Rules in the following way:

(10) $$\begin{bmatrix} \text{I stress} \\ V \end{bmatrix} \rightarrow [\text{I stress}] \; / \; \begin{pmatrix} \text{---} \ldots \overset{1}{V} \ldots]_{NAV} \\ \overset{1}{V} \ldots \text{---} \ldots] \end{pmatrix}$$

(a) COMPOUND RULE

(b) NUCLEAR STRESS RULE

In rule (10), we suppress the variable α. We interpret the rule as a sequence of two rules in accordance with the following quite general convention: a rule of the form (11) is an abbreviation for a sequence of rules of the form (12).

(11) $$X \rightarrow Y \; / \; \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{pmatrix}$$

(12) $$\begin{aligned} X &\rightarrow Y \; / \; Z_1 \\ X &\rightarrow Y \; / \; Z_2 \\ &\vdots \\ X &\rightarrow Y \; / \; Z_n \end{aligned}$$

The i^{th} rule of (12) is interpreted as stating that any symbol meeting the condition X acquires the features listed as Y when it is in a context meeting the condition Z_i. In accordance with these conventions, which will be generalized as we proceed, the rules (10a)-(10b) have precisely the same content as the sequence (8)-(9).

The rules so far discussed illustrate two general observations that have proven valid in every careful study of phonological processes that has so far been undertaken within the framework of generative grammar, namely, the following:

(13) It is always possible to order the rules in a sequence and to adhere strictly to this ordering in constructing derivations without any loss of generality as compared to an unordered set of rules or a set ordered on a different principle.

(14) Such linear ordering makes it possible to formulate grammatical processes that would otherwise not be expressible with comparable generality.[4]

[4] We shall see later that certain qualifications are necessary in the formulation of (13) and (14).

The observations (13), (14) are implicit in Bloomfield's "Menomini Morphophonemics" (1939). In Bever (1967), it is shown that the depth of ordering of Bloomfield's grammatical description is at least eleven; that is, from the linear sequence of rules constituting this grammar, a subsequence of eleven rules can be extracted with the property that the grammar becomes more complex if any two successive rules of this subsequence are interchanged in the ordering. In this same sense of depth of ordering, a depth of at least twenty-five is demonstrated in Chomsky (1951).

Neither of these statements is a necessary truth;[5] each represents an interesting and, for the present, reasonably well-confirmed empirical hypothesis. With the modification already stated as the principle of the transformational cycle, we will accept the empirical hypothesis

[5] This fact is sometimes overlooked in the case of (13). To illustrate the empirical character of (13), consider three hypothetical languages L_1, L_2, L_3, each containing the phonological segments *A*, *B*, *X*, *Y* and the lexical entries *ABY*, *BAX*. Suppose, furthermore, that in each of these languages it is a fact that *B* is realized as *X* before *Y* and that *A* is realized as *Y* before *X*. Thus the grammars contain the rules (α) and (β) as the most general statement of the facts:

(α) $B \rightarrow X \; / \; \text{---} \; Y$

(β) $A \rightarrow Y \; / \; \text{---} \; X$

Suppose now that the lexical entries *ABY* and *BAX* are realized phonetically in the following ways in L_1, L_2, L_3:

In L_1	*ABY* is realized as *YXY*	*BAX* is realized as *BYX*
In L_2	*ABY* is realized as *AXY*	*BAX* is realized as *XYX*
In L_3	*ABY* is realized as *AXY*	*BAX* is realized as *BYX*

The facts of L_1 and L_2 can be accounted for by letting the rules (α) and (β) apply in different orders: in L_1 (α) precedes (β); in L_2 (β) precedes (α). Then in L_1 we will have the derivations of (I) for the lexical entries *ABY* and *BAX*, and in L_2 we will have the derivations of (II) for the same lexical entries:

(I)

ABY	*BAX*	
AXY		BY RULE (α)
YXY	*BYX*	BY RULE (β)

(II)

ABY	*BAX*	
	BYX	BY RULE (β)
AXY	*XYX*	BY RULE (α)

Hence the hypothetical languages L_1 and L_2 support the empirical generalizations (13) and (14). However, the facts of L_3 cannot be accounted for in this fashion. As we have just seen, neither the ordering (α), (β) nor the ordering (β), (α) will give the result required, namely, that *ABY* is realized as *AXY* and that *BAX* is realized as *BYX*. Nevertheless, rules (α) and (β) state the facts in the simplest and most general way. Therefore the hypothetical language L_3 refutes the empirical hypothesis (13). In fact, L_3 supports a different empirical hypothesis concerning rule ordering, namely, that rules be unordered and that they apply simultaneously, so that each derivation has only two steps. With this convention (call it the "simultaneous application" convention), we have the derivations (III), as required for L_3:

(III)

ABY		*BAX*	
AXY	BY RULE (α)	*BYX*	BY RULE (β)

The simultaneous application hypothesis was first made explicit by Z. S. Harris (1951, Appendix to §14.32), in a discussion of an example from Bloomfield (1939) in which statement (13) was explicitly assumed. It has since been restated several times by Lamb (1964 and elsewhere), who, however, introduced a new element into the discussion by his assumption that the simultaneous application hypothesis is simpler, in some absolute sense, than the hypothesis that rules apply in sequence, in a fixed order. We see no justification for such assumptions about an absolute sense of "simplicity," in this case, nor any relevance to such assumptions if they can be given sense. The issue seems to us an empirical one; that is, the issue is whether the case posited in the hypothetical language L_3 actually is representative of natural language. So far as we know, it is not. On the contrary, the empirical evidence in natural language rules against the hypothetical situation of L_3, and therefore against the simultaneous application hypothesis and in favor of the hypotheses (13), (14). We shall have more to say about this matter as we proceed. In fact, we shall note that there are situations, formally well defined, in which something like the simultaneous application hypothesis is correct, e.g., in the case of rules that switch values of a feature. (See Chapter Eight, Sections 3, 4.) Thus the situation is complex, but, we think, quite clear.

For further discussion, see Chomsky (1964, §4.2; 1967) and Chomsky and Halle (1965).

that the rules are linearly ordered as the basis for the work to be presented here, and will give many examples that support this hypothesis. We assume, then, the following principles:

(15) (a) The rules of the phonological component are linearly ordered in a sequence $R_1, \ldots, R_n$.
(b) Each rule applies to a maximal string containing no internal brackets.
(c) After applying the rule R_n, we proceed to the rule R_1.
(d) Unless an application of R_n intervenes, the rule R_j cannot be applied after the rule R_i $(j < i)$ has applied.
(e) R_n is the rule: erase innermost brackets.

The joint effect of these principles is that the rules apply in a linear sequence to a minimal phrase of the surface structure, then reapply in the same sequence to the next larger phrase of the surface structure, and so on. When we speak of the principle of the "transformational cycle," we are referring to the empirical hypothesis (15). The statement of principle (15) is not yet sufficiently precise to resolve all questions as to how rules apply, and we will sharpen and refine it as we proceed.

In the technical terminology of the theory of generative grammar, the term "grammatical transformation" refers to a rule that applies to a string of symbols by virtue of some categorial representation of this string. We use the term "transformational" in referring to the principle just established since the rules in the cycle are transformational in the usual sense; that is, the domain of their applicability and the manner in which they apply is determined by the phrase structure of a string, not just by the sequence of elementary symbols of which the string is constituted.[6] More specifically, the application of the cyclical rules depends not only upon the formatives in the surface structure but also upon the way they are categorized. For example, the specification of N, A, or V is necessary for determining the applicability of the Compound Rule.

Notice, once again, that the principle of the transformational cycle is a very natural one. What it asserts, intuitively, is that the form of a complex expression is determined by a fixed set of processes that take account of the form of its parts. This is precisely what one would expect of an interpretive principle that applies to phrase markers, in this case, surface structures.[7]

Returning now to actual examples, let us consider the more complex phrases *black board-eraser* ("board eraser that is black"), *blackboard eraser* ("eraser for a blackboard"), and *black board eraser* ("eraser of a black board"), with the stress contours 213, 132, and 312, respectively.[8] Application of the rules discussed to the surface structure of these forms

[6] The rules involved here are, however, transformations of a very narrow and restricted class, the class referred to as "local transformations" in Chomsky (1965).

[7] Observe that the interpretive semantic rules must apply in accordance with essentially the same principle as the one stated here for the phonological rules, as has been pointed out by Fodor and Katz (1963) and by Katz and Postal (1964). The basic semantic rules apply to deep structures rather than to surface structures, however. In a sense the transformational syntactic rules also meet a similar cyclic condition. See Chomsky (1965, Chapter 3) for discussion

[8] Phoneticians might vary slightly in their description of the contours for these phrases. Whether these discrepancies are a matter of fact or of convention is a question to which we will return below. In any event, the matter is of little importance for the present. Our rules could be slightly modified to accommodate different decisions. For example, a slight revision of rule (7) would provide the contour 313 instead of 312 for the last example.

gives us the following derivations (with all occurrences of # suppressed):

(16) (a) $[_{NP}\ [_A black]_A\ [_N\ [_N board]_N\ [_N eraser]_N\ ]_N\ ]_{NP}$

1	1			RULE (2)
		1		
	1	2		RULE (10a)
2	1	3		RULE (10b)

(b) $[_N\ [_N\ [_A black]_A\ [_N board]_N\ ]_N\ [_N eraser]_N\ ]_N$

1	1			RULE (2)
		1		
1	2			RULE (10a)
1	3	2		RULE (10a)

(c) $[_N\ [_{NP}\ [_A black]_A\ [_N board]_N\ ]_{NP}\ [_N eraser]_N\ ]_N$

1	1			RULE (2)
		1		
2	1			RULE (10b)
3	1	2		RULE (10a)

These derivations illustrate the expository conventions that we will use henceforth. Let us now consider them in detail.

In the case of (16a), in the first cycle primary stress is placed on the minimal phrases *black* and *board*, which are monosyllables and therefore subject to rule (2). Also in the first cycle, primary stress is placed on *eraser* by a rule which we have not yet presented. Innermost brackets are then erased, and we return to the first of the linear sequence of transformational rules. The string now under consideration is (17), this being the only maximal string of (16a) which, at this point in the derivation, contains no internal brackets.

(17) $[_N \overset{1}{board}\ \overset{1}{eraser}]_N$

Rule (10a), the Compound Rule, is applicable to (17), and assigns primary stress on the first word, giving the stress contour 12 for this string by the conventions established previously. Since (10b) is inapplicable, we conclude this cycle, erasing innermost brackets. The string now under consideration is

(18) $[_{NP} \overset{1}{black}\ \overset{1}{board}\ \overset{2}{eraser}]_{NP}$

Rule (10a) is not applicable to this string, so we turn to rule (10b), the Nuclear Stress Rule, which assigns primary stress to *board*, weakening all other stresses in (18) by one. This gives the contour 213 as the final line of derivation (16a).

The derivation (16b) has the same first cycle as (16a), but for the second cycle, the string under consideration is the noun *blackboard* rather than the noun *board-eraser*. The

Compound Rule assigns to this noun the contour 12. Erasing innermost brackets, we proceed to the next cycle, considering now the noun *blackboard eraser* (whereas at the analogous stage of the derivation (16a), we considered the noun phrase *black board-eraser*). Being a noun, this string is subject to the Compound Rule, so that primary stress is placed on the first word, giving the contour 132.

Now consider the derivation (16c). The first cycle is exactly the same as in the other two derivations of (16). But in the second cycle we consider not the noun *board-eraser*, as in (16a), nor the noun *blackboard*, as in (16b), but the noun phrase *black board*, meaning "board that is black." To this, only the Nuclear Stress Rule applies, placing primary stress on the second word. This concludes the second cycle. In the third cycle we consider the noun *black board eraser*, which at this stage has the contour 211. The Compound Rule (10a) applies to this string, assigning primary stress to the leftmost primary-stressed vowel and weakening all the others. This gives the desired contour 312.[9]

To illustrate the transformational cycle with one more example, consider the noun phrase *John's blackboard eraser*, which undergoes the following derivation (where D stands for the category "determiner"):

(19)

$$[_{NP}\ [_{D}\mathit{John's}]_{D}\ [_{N}\ [_{N}\ [_{A}\mathit{black}]_{A}\ [_{N}\mathit{board}]_{N}\ ]_{N}\ [_{N}\mathit{eraser}]_{N}\ ]_{N}\ ]_{NP}$$

1 1 1

RULE (2): 1

RULE (10a): 1 2

RULE (10a): 1 3 2

RULE (10b): 2 1 4 3

The phrase *blackboard eraser* undergoes the three-cycle derivation (16b); the determiner *John's* receives its primary stress on the first cycle. In the fourth cycle, the string under consideration is the noun phrase *John's blackboard eraser*, with the stress contour 1132. The Nuclear Stress Rule assigns primary stress to the rightmost primary-stressed vowel, weakening all others, and giving the stress contour 2143.[10]

Suppose that the phrase *John's blackboard eraser* appears in the context —— *was stolen*. The whole phrase in this case is a sentence, i.e., is bounded by $[_{S}\ \ldots\ ]_{S}$. The word *stolen* will receive primary stress on the first cycle, and *John's blackboard eraser* will have the derivation (19). In the final cycle, at the level of $[_{S}\ \ldots\ ]_{S}$, primary stress will be placed on *stolen* by the Nuclear Stress Rule, giving *John's blackboard eraser was stolen*.[11]

[9] Though examples (16a) and (16c) may appear somewhat artificial, the reality of the syntactic patterns they illustrate can hardly be doubted. They appear, for example, in such phrases as *American history-teacher*, meaning "American teacher of history," which is analogous to (16a) and has the same stress contour 213; and in *American-history teacher* meaning "teacher of American history," which is analogous to (16c) and also has the stress contour 312 (or 313—see note 8). We assume here that the word *American* receives primary stress on the second syllable, although we have not yet given the rules that determine this. Similarly, the phrases *civil rights bill* and *excess profits tax* are of the form illustrated in (16c), whereas *uncivil game warden* or *excessive profits tax* are of the form illustrated in (16a). There are many other exact or near minimal pairs, e.g., *civil engineering student* ("student of civil engineering" or "polite student of engineering"), *small boys school* ("school for small boys" or "boys' school that is small").

[10] See note 8.

[11] See note 8. To prevent *was* from receiving primary stress by rule (2), we restrict this rule, as a first approximation, to the lexical categories, namely, noun, adjective, verb. We assume, on syntactic grounds, that the auxiliary *be* is not introduced as a member of a lexical category.

Suppose that the phrase *John's blackboard eraser* appears in the context *take*——, the whole constituting a sentence (in this case, an imperative). The word *take* receives primary stress and *John's blackboard eraser* receives the contour 2143 by the derivation (19). In the final stage of the cycle, the Nuclear Stress Rule (10b) places primary stress on *black*, giving the final contour 23154.

These examples show how complex and varied phonetic representations are determined by very simple rules when the principle of the transformational cycle is presupposed; in other words, they illustrate the kind of evidence that can be offered in support of the hypothesis that it is the principle of the transformational cycle that underlies the phonetic interpretation of utterances. Observe that no rules at all are needed beyond those required for the most elementary phrases. The interplay of these rules in more complex phrases is determined by the principle of the transformational cycle, which is, it should be noted, not a rule of English grammar but rather a general principle governing the applicability of phonological rules in any grammar.

Notice that the rules, as presented, assign a different internal stress contour to the phrase *John's blackboard eraser* depending upon whether it appears in subject or object position in the surface structure. In subject position, as in the context —— *was stolen*, the contour of the phrase is 3254, with the same internal relations of stress as in the phrase in isolation, though weakened in each case by one degree. In object position, on the other hand, as in the context *take*——, the contour of the phrase is 3154, with internal relations that are different from those of the phrase in isolation. Similarly, a simple adjective-noun construction such as *sad plight* will have the contour 21 in isolation, the contour 32 in the context *his*—— *shocked us*, and the contour 31, with different internal relations, in *consider his*——. As the structure of the sentence becomes more complex, the internal relations of stress within a phrase of this sort will continually be modified. Thus in the sentence *my friend can't help being shocked at anyone who would fail to consider his sad plight*, the surface structure might indicate that the word *plight* terminates no less than seven phrases to which the Nuclear Stress Rule applies, so that successive applications of this rule would give the contour *sad plight*. Presumably, the actual internal relations of stress in *sad plight* are the same, in this case, as in *consider his sad plight*, or even in *sad plight* in isolation.

In connection with this problem, several comments are called for. First, it is very likely that certain readjustment rules of the sort mentioned in Chapter One, page 10, must be applied to surface structures before the application of phonological rules, deleting structure and restricting the number of applications of the transformational cycle (and, consequently, the fineness of stress differentiation). Second, it is necessary to formulate a principle for interpretation of phonetic representations that nullifies distinctions that go beyond a certain degree of refinement. Third, there may very well be additional principles that modify the convention weakening stress when primary stress is placed in a complex construction. Finally, it is necessary to take note of the qualifications with respect to phonetic representation in general that we discuss in the next section.

Before leaving the topic of stress contours within phrases, we should make it quite clear that the rules discussed above give accurate results only for very simple constructions. We have not investigated the problem of determining the stress contours of complex phrases of varying syntactic types; our investigation has been limited to the very restricted types of constructions that have been discussed in the literature on English phonetics and phonology of the past several decades. There is, for the moment, little useful data on more complex

constructions. Such observations as have been made suggest that the problem of extending this description to a wider class of cases may be nontrivial. For example, Stanley Newman, in his important article on English intonation (1946), points out that in the sentence *he has plans to leave*, the contour on *plans to leave* is rising if the meaning is, roughly, "he intends to leave," but is falling if the meaning is "he has documents to leave." It is not at all clear what features of syntactic structure determine this difference. Another class of phenomena not accounted for are those involving obligatory contrastive stress (sometimes stress shift) as determined by syntactic parallelism, as in such sentences as *he wanted to study electrical rather than civil engineering*, or *instead of encouraging the teacher to make the work interesting, the school administrators actually discourage her*. Many other problems can be cited, all indicating that many questions of fact and, perhaps, of principle still remain unresolved in this area.

2. *On the reality of phonetic representation*

Utilizing the principle of the transformational cycle, the speaker of English can determine the phonetic shape of an utterance on the basis of such rules as the Compound and Nuclear Stress Rules, even though the particular utterance may be quite new to him. He need not deal with the stress contour as a property of the utterance independent, in whole or in part, of its syntactic organization. There is no doubt that stress contours and many other phonetic properties are determined for new utterances with quite a bit of consistency among speakers. This is a fact that must be accounted for by an empirically adequate grammar. In the case of English we can approach an explanation by incorporating in the grammar such rules as the Compound and Nuclear Stress Rules and by postulating the principle of the transformational cycle. Before going on to investigate the rules of English in greater detail, let us briefly consider the question of how these rules and the general principles that govern their applicability relate to psychological processes and to physical fact.

We might suppose, on the basis of what has been suggested so far, that a correct description of perceptual processes would be something like this. The hearer makes use of certain cues and certain expectations to determine the syntactic structure and semantic content of an utterance. Given a hypothesis as to its syntactic structure—in particular its surface structure—he uses the phonological principles that he controls to determine a phonetic shape. The hypothesis will then be accepted if it is not too radically at variance with the acoustic material, where the range of permitted discrepancy may vary widely with conditions and many individual factors. Given acceptance of such a hypothesis, what the hearer "hears" is what is internally generated by the rules. That is, he will "hear" the phonetic shape determined by the postulated syntactic structure and the internalized rules.

Among the internalized rules are some that are particular to the language in question and thus must have been learned; there are others that simply play a role in setting the conditions on the content of linguistic experience. In the present case, it would be reasonable to suggest that the Compound and Nuclear Stress Rules are learned, while the principle of the transformational cycle, being well beyond the bounds of any conceivable method of "learning," is one of the conditions, intrinsic to the language-acquisition system, that determines the form of the language acquired. If this assumption is correct, we would expect the principle of the transformational cycle to be a linguistic universal, that is, to be consistent

with the empirical facts for all human languages;[12] the Compound and Nuclear Stress Rules, on the other hand, might be in part language-specific.

We do not doubt that the stress contours and other phonetic facts that are recorded by careful phoneticians and that we will study here constitute some sort of perceptual reality for those who know the language in question. In fact we are suggesting a principled explanation for this conclusion. A person who knows the language should "hear" the predicted phonetic shapes. In particular, the careful and sophisticated impressionistic phonetician who knows the language should be able to bring this perceptual reality to the level of awareness, and there is ample evidence that phoneticians are capable of doing this. We take for granted, then, that phonetic representations describe a perceptual reality. Our problem is to provide an explanation for this fact. Notice, however, that there is nothing to suggest that these phonetic representations also describe a physical or acoustic reality in any detail. For example, there is little reason to suppose that the perceived stress contour must represent some physical property of the utterance in a point-by-point fashion; a speaker who utilizes the principle of the transformational cycle and the Compound and Nuclear Stress Rules should "hear" the stress contour of the utterance that he perceives and understands, whether or not it is physically present in any detail. In fact, there is no evidence from experimental phonetics to suggest that these contours are actually present as physical properties of utterances in anything like the detail with which they are perceived. Accordingly, there seems to be no reason to suppose that a well-trained phonetician could detect such contours with any reliability or precision in a language that he does not know, a language for which he cannot determine the surface structure of utterances.

Considerations of this sort lead us to suspect that the question of how highly differentiated the stress contours in a representation should be is of little significance. In a complex utterance with a rich surface structure, the rules outlined in the preceding section will lead to a stress contour of many levels. There may be no empirical sense to the question of whether the resulting representation is correct in full detail. Because of the completely impressionistic character of judgments of relative stress, decisions over a broad range are of little value. It is not at all surprising that there should be great difficulty, within impressionistic phonetics, in determining how many stress levels should be marked and how they are distributed in utterances that exceed a certain degree of complexity. The shape and the degree of differentiation of a stress contour are largely determined by obligatory rules and are therefore below the level of systematically significant representation. Once the speaker has selected a sentence with a particular syntactic structure and certain lexical items (largely or completely unmarked for stress, as we shall see), the choice of stress contour is not a matter subject to further independent decision.[13] That is, he need not make a

[12] In one sense, a general principle counts as a linguistic universal if it is compatible with the facts for all human languages. As linguists, of course, we are concerned not with principles that happen by accident to be universal in this sense, but rather with those that are universal in the domain of all possible human languages, that is, those that are in effect preconditions for the acquisition of language. (See the discussion in Chapter One, p. 4.) Such principles, and such alone, can serve to explain and account for the phenomena of particular languages. The distinction in question is not easy to draw, but is no less crucial for this reason.

Notice, incidentally, that the transformational cycle might apply vacuously in a certain language, in particular if the language has very shallow surface structure. Thus a highly agglutinative language might be expected to offer little or no support for the principle of the transformational cycle, at least within the bounds of a word. This, if true, would be entirely irrelevant to the status of this principle as a linguistic universal.

[13] We assume that the position of emphatic stress is marked in the surface structure, and we neglect matters that we have assigned to the theory of performance (see Chapter One, p. 3).

choice among various "stress phonemes" or select one or another "superfix." With marginal exceptions, the choice of these is as completely determined as, for example, the degree of aspiration. Similarly, a hearer who has grasped the structure and morphemic constitution of an utterance from a rough sampling of the physical input need not attend to stress variation, to whatever extent this may actually be a physical property of utterances.

It is to be expected that determined phonetic features should be quite difficult for the user of the language to learn to identify, whether they involve stress or degree of aspiration (where undoubtedly there are many levels, predictable, at least roughly, by general rules).[14] The apparent ease with which phoneticians trained in the same conventions can, to a large extent, agree on the assignment of four or five stresses in utterances may very well be traceable to their ability, as speakers of the language, to grasp the syntactic structure of utterances and to assign to them an "ideal" stress contour by the rules of the transformational cycle. Such an achievement may have little to do with any physical fact. This is, incidentally, a matter which should be subject to experimental investigation.[15]

To summarize this discussion of phonetic representation, we do not doubt that representations of stress contours and similar predictable phenomena correspond, up to a point, to some perceptual reality that can be brought to consciousness with training and care. That this must be true is shown by the fact that phoneticians trained in the same system of conventions can reach considerable agreement in transcribing novel utterances in languages that they know. These perceptual facts may be of interest only to the extent that they provide data for testing empirical hypotheses such as the principle of the transformational cycle. Accordingly, perceived stress contours are of very great linguistic interest since they offer evidence bearing on this hypothesis, whereas degree of aspiration will be of no linguistic interest if, as one might suspect, it is determined by principles of little depth or generality. Furthermore, the representation of the perceptual facts is likely to be governed in part by arbitrary convention or irrelevant cognitive limitations after a certain degree of complexity is reached. Thus, it is impossible to expect (and, for purposes of investigating linguistic structures, unnecessary to attain) a complete correspondence between the records of the impressionistic phonetician and what is predicted by a systematic theory that seeks to account for the perceptual facts that underlie these records.

3. *The transformational cycle within the word*

Let us return now to the problem of how the phonological component of a grammar is organized, and the more specific matter of the rules of English phonology. In the derivations given in Section 1, we did not provide rules for determining stress placement in the word *eraser* or, for that matter, in any word that is not a monosyllable (see rule (2)). In fact, it is evident that *eraser* is itself a complex form based on the verb *erase* and an agentive

[14] As noted, there is no acoustic evidence to support the view that perceived stress contours correspond to a physically definable property of utterances. However, even if such differentiations did exist along a single dimension of the acoustic signal, there would be some reason to doubt that they might be identified by phoneticians. There is evidence that even under experimental conditions, where complex stimuli are to be sorted along several dimensions, more than two or three distinctions along each dimension will overload the perceptual capacity. See Pollack and Ficks (1954) and Miller (1956).

[15] P. Lieberman (1965) has shown that a phonetician who is capable of describing a pitch contour with great accuracy in isolation may represent this very same contour quite differently when it is associated with an utterance of his language. This strongly suggests that what the phonetician "hears" in utterances depends very heavily on internalized rules that predict perceived phonetic shape. Similar results were obtained for stress.

affix. Thus, at the level where phonological rules of the kind we are now considering become applicable, the structure of this item is something like (20):[16]

(20) $$[_N\# [_V\# erase \#]_V r \#]_N$$

If the principle of the transformational cycle is perfectly general, then this word too should have more than one cycle in its derivation. The rules should first apply to the underlying verb *erase* and then, in the next cycle, to the noun *eraser*. The verb *erase* is bisyllabic, and we see that stress is placed on the second syllable. As a first approximation to the rule of stress placement for lexical items, we can formulate the rule (21), which places primary stress on the final vowel of the string under consideration where this item is a noun, adjective, or verb. The symbol C_0, as before, stands for a string of zero or more consonants.

(21) $$V \rightarrow [1 \text{ stress}] \ / \ X \text{ —— } C_0]_{NAV}$$

Notice that rule (21) now includes, as a special case, rule (2), which placed primary stress on the only, hence final, vowel of a monosyllabic item. We can thus dispense with rule (2), and the rules of stress placement become rules (21), (10a) and (10b) (the Compound and Nuclear Stress Rules), and (7), which appears to be quite marginal.

There is a difficulty, however. If these rules apply in a cycle, rule (21) will be applicable to nouns such as *blackboard*, *blackboard eraser*, and so on, incorrectly assigning primary stress to the final vowel. We must therefore place some restriction on rule (21) to eliminate this possibility. The simplest way to do this is to require that the string to which (21) is applied must contain no occurrences of the boundary #. We therefore add to rule (21) the condition (22):

(22) *X* contains no internal occurrence of #.

With rule (21) replacing rule (2), we have provided sufficient information to complete the derivations that were given as examples of the operation of the transformational cycle. In the first stage, rule (21) applies to assign primary stress to the final vowel of each of the items *black*, *board*, *John*, *erase*. The second cycle will be vacuous in the case of *John's* or *eraser*, stress simply being reassigned to the stressed vowel.[17] Otherwise, the derivations proceed as before.

The transformational cycle operates within word boundaries in a much more far-reaching and extensive way than suggested by examples such as these. In complex derivational forms, for example, it seems quite natural to suppose that the phonetic shape of the full form is determined by general rule from the ideal representation of its parts in much the same way as in syntactic constructions. Investigation of English and other languages confirms this expectation and permits us to formulate the principle of the transformational cycle in full generality, applying to all surface structure whether internal or external to the word. The word is, as we shall see, a significant phonological unit, but its unique properties do not lead to violation of the general principle of the transformational cycle. We assume, then, that the cycle operates from the minimal units included in (or, in special cases, constituting) words up to the maximal domain of phonological processes, with no discontinuity.

[16] On the placement of # boundaries, see Chapter One, pages 12–14.

[17] We shall see that the reason for the inapplicability of any rules in the second cycle of these forms is actually quite different from what is suggested here. In both cases it is the # boundary preceding the affix which blocks all phonological rules that would otherwise be applicable.

4. The segmental phonology of English—a first approximation

We have described the phonological component as a system of rules, organized in accordance with the principle of the transformational cycle, which maps surface structures into phonetic representations, where a surface structure is a labeled bracketing of a string of formatives. Furthermore, we have been assuming that the formatives can themselves be regarded as strings, consisting of consonants and vowels. The lexicon, which is a part of the syntactic component of the grammar, determines the intrinsic structure of a formative in terms of phonological properties: in particular, the lexicon determines how a formative is represented as a string of consonants and vowels. We will refer to the consonants and vowels that constitute a formative as its "segments." The phonological rules modify the segmental structure of a string of formatives in accordance with the specified labeled bracketing. At the termination of the transformational cycle, all labeled bracketing has been erased, and we are left with a string of phonological elements which we will also refer to as segments, in this case "phonetic segments." These segments too can be analyzed as consonants and vowels of various types. We assume that linguistic theory includes a universal phonetic alphabet—of a sort that we will later describe in detail—which provides a uniform, language-independent system for the representation of phonetic segments. In brief, then, the phonological component maps a surface structure into a string of universal phonetic segments.

Let us for the moment assume a standard phonetic system for the representation of consonants and turn our attention to the system of English vowels.

For our immediate purposes, we may regard a formative as a string of consonants and "vocalic nuclei." The vocalic nuclei may be "simple," as in the boldface positions of *pit, pet, pat, put, putt, analyze*. We will use the phonetic symbols i, e, æ, u, ʌ, ə, respectively, for these simple vocalic nuclei, delaying a more detailed analysis until later. The segment represented as ə will be referred to as the "reduced vowel."

In addition to simple vocalic nuclei, there are "complex vocalic nuclei," such as those that appear in the boldface positions in *confide, feed, fade, feud, road*, and others. For the time being, we will use the symbols *I, E, A, U, O*, respectively, for the complex nuclei of the cited forms; that is, we use each capital letter with its conventional name as its phonetic value.

Following this convention, we will have quasi-phonetic spellings such as the following:

(23)

erase	ErÁs
irate	IrÁt
mutation	mUtÁšən
ecumenical	ekUménikəl[18]
cupidity	kUpídItE
citation	sItÁšən
maintain	mAntÁn
collapse	kəlæps

[18] Or, perhaps, [ekUmenəkəl]. As indicated in the Preface, we will generally follow the phonetic representations of Kenyon and Knott, which agree quite well with our own normal speech in most respects. Although there are some differences which we will comment on later, none of them are very crucial, and for the moment we can ignore them.

The representation of other vocalic nuclei and a more detailed analysis of all of these elements will concern us in later chapters. We will discover, in fact, that the representations just proposed are somewhat more than a mere notational convenience.

In terms of the above notions, we can distinguish between "weak clusters" and "strong clusters" in the following way. A weak cluster is a string consisting of a simple vocalic nucleus followed by no more than one consonant; a strong cluster is a string consisting of either a vocalic nucleus followed by two or more consonants or a complex vocalic nucleus followed by any number of consonants. In either case, the cluster is assumed to be followed either by a vowel or by the boundary symbol # (with possible intrusions of the + boundary). These definitions will be emended and made more precise later on.

Using the symbol S for a strong cluster and W for a weak cluster, we can see that the items of (23) are phonetically of the following form in terms of clusters (with initial consonants omitted):

(24)

ErÁs	SS
IrÁt	SS
mUtÁšən	SSW
ekUménikəl	WSWWW
kUpídItE	SWWS
sItÁšən	SSW
mAntÁn	SS
kəlæps	WS

5. More on the transformational cycle within the word

We can now proceed to deepen the account of stress placement within words. Rule (21), the only rule given so far that places stress within words, assigns primary stress to the final vowel of the string under consideration. Thus it assigns primary stress to the final syllable of words such as *evade, supreme, exist, absurd*. Observe, however, that all these examples have final strong clusters phonetically. In fact, if a verb or adjective has a final weak cluster, then stress is placed on the penultimate rather than the final syllable. Thus we have words such as *relish, covet, develop, stolid, common, clandestine*, all with penultimate stress and final weak clusters.[19] These observations suggest that rule (21) should be divided into two cases, the first assigning primary stress to the vowel preceding a final weak cluster, the second assigning primary stress to the final vowel of the string under consideration. We can give this rule in the following form:

(25) $$V \rightarrow [1 \text{ stress}] \; / \; X \text{——} C_0(W)]$$

where X contains no internal occurrences of # (see condition (22)) and W is a weak cluster. We interpret (25) as an abbreviation for two rules, in accordance with the general convention that a rule of the form (26), with a string in parentheses, is an abbreviation for the

[19] Exceptions to the rules we are now sketching will readily come to mind. To a considerable extent they will be taken care of by the more careful formulation given in the next chapter. Exceptions do remain, however. (See the Preface on the subject of exceptions.)

Notice that the rule we are discussing here is, in effect, the familiar Latin stress rule.

sequence of rules (27) (where either Z or Q contains ——):

(26) $$X \rightarrow Y \;/\; Z(P)Q$$

(27) (a) $X \rightarrow Y \;/\; ZPQ$
(b) $X \rightarrow Y \;/\; ZQ$

The order in (27) is crucial: in a sequence of rules abbreviated by the parenthesis notation, as in (26), the case (27a) that includes the string in parentheses is applicable before the case (27b) without the parenthesized string. In accordance with these conventions, rule (25) is an abbreviation for the two rules (28a) and (28b), in that order:

(28) (a) V → [1 stress] / X —— C_0W]
(b) V → [1 stress] / X —— C_0]

Words such as *relish*, *develop*, *common*, with final weak clusters, are subject to (28a) and receive penultimate stress. Words such as *evade*, *supreme*, *exist*, with final strong clusters, are not subject to (28a) and receive stress on the final syllable by (28b).

There is one additional condition to be noted in connection with rule (25). Suppose that we apply this rule to a word with a final weak cluster, such as *edit*. By case (28a), primary stress is placed on the penultimate syllable, giving *edit*. But then, by case (28b), primary stress will be shifted to the final syllable and the first syllable will be weakened to [2 stress], resulting in the incorrect form **edit*. The simplest and most general way to avoid this is to establish a condition on the parenthesis convention itself. In fact, in all descriptive work in generative grammar with which we are familiar, it has been tacitly assumed that in the case of a rule such as (26), the two subcases (27a) and (27b) are ordered not only as shown, but are "disjunctively ordered," in the sense that if rule (27a) applies, then rule (27b) is not permitted to apply. Thus a sequence of rules abbreviated in terms of the parenthesis notation constitutes a disjunctively ordered block; as soon as one of these rules is applied, the remaining rules are skipped within any one cycle of a derivation. We now establish this as a general convention with regard to the parenthesis notation, to be extended and generalized as we proceed. We thus extend the general theory of the organization of a grammar expressed in the principle of the transformational cycle, by observing that certain subsequences of the linearly ordered rules may be disjunctively ordered. To return to the rules we have been discussing, the two cases (28a) and (28b) abbreviated by (25) will be disjunctively ordered, and the difficulty noted at the beginning of this paragraph will not arise; once case (28a) has applied to give the correct form *edit*, then case (28b) is prevented, by the principle of disjunctive ordering, from applying to that form.

Like other general conditions on the organization of a grammar, the convention just proposed constitutes an empirical hypothesis subject to refutation by linguistic fact. The hypothesis is, in this case, that if a sequence of rules is to be abbreviated by the parenthesis convention,[20] then this sequence forms a disjunctively ordered block. Obviously, this is not a necessary truth, by any means.

[20] The question of when a sequence of rules is to be abbreviated by the parenthesis convention is not a matter of choice but rather one of fact. That is, the convention regarding parentheses is just one part of an evaluation procedure to be applied to grammars. This procedure is perfectly general (language-independent) and performs the function of determining which of the grammars consistent with the data is to be selected as the grammar of the language for which the data provide a sample. For discussion, see Chomsky (1965) and many earlier references.

The matter of defining "optimal representation" is nontrivial. In the ensuing discussion we make certain tacit assumptions about "optimality" that will be explored further in Chapter Three, Section 1. See Chomsky (1967) for further discussion.

It is not to be expected that an absolutely crucial test case for this hypothesis will be very easy to come by. In any real case, there will presumably be other aspects of a grammatical description which, if modified, will allow this hypothesis to be retained in the face of superficially disconfirming evidence. This is the usual situation when an empirical hypothesis of such generality is at issue. Still, it is quite clear what sort of evidence is relevant to increasing or diminishing the plausibility of the hypothesis.

Returning now to the problem of stress assignment, we see at once that rule (25) requires refinement and elaboration if it is to account for the facts. Each of the examples given to illustrate the rule contains just a single formative. Where a word has an internal analysis in terms of formatives, rule (25) must apply in a slightly different way. To see this, consider the derived forms *person*+*al*, *theatr*+*ic*+*al*, *anecdot*+*al*, *dialect*+*al*. If rule (25) were to apply directly to these forms, it would assign primary stress to the penultimate syllable (the final cluster *-al* being weak), giving **personal*, **theatrical*, *anecdotal*, *dialectal*, only the last two of which are correct. Notice that all four words would be assigned primary stress in the correct way by rule (25) if the affix *-al* were excluded from consideration at the point when the rule is applied. The residual forms *person-* and *theatric-*, with final weak clusters, would have primary stress assigned to their penultimate syllables by case (28a); the forms *anecdOt-* and *dialect-*, on the other hand, would be exempt from (28a) because of their strong final clusters and would instead have primary stress assigned to the final syllable by case (28b). This observation is in fact quite general for affixes, and we therefore replace rule (25) by the following sequence of rules:

(29) (a) V → [1 stress] / X —— C_0(W)+affix]
(b) V → [1 stress] / X —— C_0(W)]

Clearly there is a generalization being missed by the formulation (29), for the obvious similarity between the two cases is not expressed. To permit us to capture generalizations of this sort, we extend our notations to permit rules such as (30):

(30) $$X \rightarrow Y \;/\; Z \text{——} R \;/\; P \text{——} Q$$

In general, a rule of the form (31) can be regarded as an abbreviation for the rule (32), where Z and R are strings:[21]

(31) $$X \rightarrow Y \;/\; Z \text{——} R$$

(32) $$ZXR \rightarrow ZYR$$

Following this convention, we interpret (30) as an abbreviation for (33), where Z and R are strings:

(33) $$ZXR \rightarrow ZYR \;/\; P \text{——} Q$$

This is now a rule of a familiar form. Reapplying the convention that defines (31) in terms of (32), we interpret (33) as an abbreviation for (34):

(34) $$PZXRQ \rightarrow PZYRQ$$

[21] We will give more precise definitions of these notions in Chapter Eight. For the present, one can think of rule (31) (equivalently, (32)) as stating that a linguistic element of the form X is extended to contain the features Y (or is modified to contain Y, if Y differs in some respect from X) when this element of the form X appears in a context of the form Z —— R. There are ambiguities in this account; they will be resolved later, and are not of the sort that should lead to misunderstanding in the present context.

Thus, when Z and R are strings, the notation (30) is well-defined. Suppose, however, that Z and R are not strings, but notations of any complexity, including braces, parentheses, and so on. Then it would not do to say that (31) is an abbreviation for (32); rather, (31) is an abbreviation for the sequence of rules (35), determined by the conventions for braces, parentheses, etc. The sequence (35) is then an abbreviation for the sequence (36), by the convention just stated.

(35)
$$\begin{pmatrix} X \rightarrow Y \ / \ Z_1 \text{——} R_1 \\ X \rightarrow Y \ / \ Z_2 \text{——} R_2 \\ \vdots \\ X \rightarrow Y \ / \ Z_m \text{——} R_m \end{pmatrix}$$

(36)
$$\begin{pmatrix} Z_1XR_1 \rightarrow Z_1YR_1 \\ Z_2XR_2 \rightarrow Z_2YR_2 \\ \vdots \\ Z_mXR_m \rightarrow Z_mYR_m \end{pmatrix}$$

This leaves us with only the problem of explaining the meaning of (30) in the case when Z and R involve notations such as braces and parentheses. Since (31), in this case, is an abbreviation for (35) (ultimately, (36)), the conventions already given will interpret (30) as an abbreviation for (37):

(37)
$$\begin{pmatrix} X \rightarrow Y \ / \ Z_1 \text{——} R_1 \\ X \rightarrow Y \ / \ Z_2 \text{——} R_2 \\ \vdots \\ X \rightarrow Y \ / \ Z_m \text{——} R_m \end{pmatrix} / \ P \text{——} Q$$

The above can be seen to be (35) (or, equivalently, (36)) in the context P——Q. By the usual brace conventions, we can now interpret (37) as an abbreviation for (38):

(38)
$$\begin{pmatrix} X \rightarrow Y \ / \ Z_1 \text{——} R_1 \ / \ P \text{——} Q \\ X \rightarrow Y \ / \ Z_2 \text{——} R_2 \ / \ P \text{——} Q \\ \vdots \\ X \rightarrow Y \ / \ Z_m \text{——} R_m \ / \ P \text{——} Q \end{pmatrix}$$

In (38), each Z_i and R_i is a string of symbols, so that (38) is itself interpretable by the convention that gives (30) as an abbreviation for (33).

We see, then, that there is a very natural way of interpreting familiar conventions so that a rule of the form (30) has, in effect, the following intuitive meaning: first, expand the context P——Q, in accordance with the brace and parenthesis conventions, into the sequence of its special cases P_1——$Q_1, \ldots, P_k$——Q_k; next, apply the rules abbreviated as $X \rightarrow Y \ / \ Z$——R in the usual sequence, under the condition that the element ZXR under consideration is in the context P_i——Q_i; next, apply the same rules under the condition that the element ZXR is in the context P_2——Q_2; etc.

With these notational remarks, we can return to the generalization left unexpressed in rule (29) which can now be captured by the following rule:

(39)
$$\text{V} \rightarrow [1 \text{ stress}] \ / \ X \text{——} C_0(\text{W}) \ / \ \text{——}(+\text{affix})]$$

where W is a weak cluster, C_0 is a string of zero or more consonants, and X does not contain an internal # boundary.[22] Our conventions interpret (39) as an abbreviation for

[22] Actually, the affix must be restricted to a glide or to a monosyllabic formative with a simple vocalic nucleus, for reasons that will be developed in the next chapter.

the following sequence of rules:

(40)
(a) V → [1 stress] / X——C_0W+affix]
(b) V → [1 stress] / X——C_0+affix]
(c) V → [1 stress] / X——C_0W]
(d) V → [1 stress] / X——C_0]

The parenthesis convention proposed earlier imposes the following ordering conditions on (40a–d): (1) the order of application is (a), (b), (c), (d), as given; (2) if case (a) applies, then case (b) is inapplicable; (3) if case (c) applies, then case (d) is inapplicable; (4) if either case (a) or case (b) applies, then cases (c) and (d) are inapplicable. Summarizing, the convention implies that the ordering of (40) is totally disjunctive; if one case applies, then all later cases are skipped.

In forms such as *person+al* and *theatr+ic+al*, case (a) of rule (40) assigns primary stress in the antepenultimate syllable. Case (b) of (40) applies to words such as *dialect+al* and *anecdOt+al*, assigning primary stress in the penultimate position, which contains a strong cluster. Cases (c) and (d) are simply the two cases of rule (25); they apply to such words as *edit* and *develop*, assigning penultimate stress, and to words such as *evade* and *supreme*, assigning primary stress in the final syllable. Rule (39) thus expresses in a precise way the linguistically significant generalization that underlies this class of examples.

Notice that some of these examples involve more than one cycle. The word *theatrical*, for example, is clearly derived from *theater*, which will receive primary stress on the initial syllable in the first cycle (by a rule which will be given in the next chapter); thus, in isolation the stress will be in that position. But in the second cycle, the stress is shifted to the second (antepenultimate) syllable by rule (39). We thus have the derivation (41). (Recall that we assume all formatives to be automatically bounded by +, by convention. We therefore need not indicate all occurrences of this boundary in a derivation.)

(41)
$[_A\ [_N theatr]_N\ ic+al]_A$
1 (RULE TO BE GIVEN)
21 RULE (39), CASE (40a)

The stress on the first syllable is then weakened as a special case of rules that we will go into later.

Suppose that we have a still more complex form such as *theatricality*, for example. For this form, the same rules provide the following derivation:[23]

(42)
$[_N\ [_A\ [_N theatr]_N\ ic+al]_A\ i+ty]_N$
1 (RULE TO BE GIVEN)
21 RULE (39), CASE (40a)
32 1 RULE (39), CASE (40a)

[23] The analysis of *-ity* as *i+ty* might be disputed, but it seems well motivated on morphological grounds. There is, first of all, a noun-forming affix *-ty* (*loyalty*, *novelty*, etc.) Furthermore, the forms in *-ity* often have other derived forms with affixes beginning with *-i* (*sanctity–sanctify–sanctitude*, *clarity–clarify*, etc.), which suggests that *-i-* is a stem-forming augment. We shall see, in fact, that there are good reasons to suppose that no affixes are polysyllabic.

As rule (39) is stated, this analysis of *-ity* is necessary. From considerations presented in the next chapter, however, it can be shown that even if *-ity* were to be analyzed as a single formative, the rules would still provide the derivation (42). Therefore, in this instance at least, phonological considerations do not require the analysis into two formatives.

There is a generally accepted convention to the effect that secondary stress appears within a word only if it is the main stress within that word. Accordingly, we add the following rule:

(43) Within a word, all non-main stresses are weakened by one.

The exact status of this rule, which we will call the Stress Adjustment Rule, is a matter to which we will return below. We will see, in fact, that it becomes a special case of the Nuclear Stress Rule (10b), when the latter is properly formulated. The Stress Adjustment Rule (43) converts *theatricality* to *theatricality*, which we can take to be the phonetic representation for this word up to the degree of detail we have discussed so far.

In the same manner, rule (39) assigns stress contours to many complex forms, in accordance with the principle of the transformational cycle. We can thus account for a substantial class of cases in a very simple and general way.

Actually, rule (39) may be extended somewhat further. Consider pairs of words such as:

(44)

phótograph	*photosýnthesis*
mónolith	*monománia*
télescope	*telekinésis*
prótoplasm	*protozóa*

Each of the forms consists of a prefix (*photo-*, *mono-*, *tele-*, *proto-*) followed by a stem (which may, in certain cases, function as an independent word). With minimal assumptions about surface structure, *photograph*, for example, will be represented $[_N photo\ [_{STEM} graph]_{STEM}\ ]_N$. In a case like *photosynthesis*, the bracketing will be the same, but *synthesis* will be labeled as a noun rather than a stem.

We note that primary stress falls on the prefix if the stem is monosyllabic,[24] and on the stem if the stem is polysyllabic. Though this observation will be modified slightly when a larger class of cases is considered, it can be accepted as a first approximation. We notice further that stress placement on the prefix is in accordance with rule (39); that is, by case (40c) (= (28a)), primary stress is assigned to the syllable preceding the final weak cluster of the prefix. (For reasons which appear below, the final vowel of *photo*, *mono*, etc., is lexically lax though in some positions it is phonetically tense.)

Using these observations and the assumed surface structure, we can account for the forms in (44) with a rule that accomplishes the following. After primary stress has been assigned to the stem (or inner noun) in the first cycle, it will be shifted left to the prefix if the stem (or inner noun) is a monosyllable, that is, if the form has a final stressed syllable when it enters the second cycle. For example, *photograph* will enter the second cycle as *photograph*, with a final stressed syllable, and our new rule will then shift the stress back to give *photograph*. The form *photosynthesis*, on the other hand, will enter the second cycle as *photosynthesis*; since the syllable that is stressed is not final, the new rule will not apply and the stress will remain on the inner noun. We can now proceed to formulate the rule as follows:

(45) $$V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0(W)\ /\ \text{——} \hat{\Sigma}]$$

[24] We are using the term "monosyllabic" in a phonological, not a phonetic, sense in this context. Thus *plasm* is phonologically monosyllabic (cf. *plasma*) but phonetically bisyllabic, since postconsonantal nasals become syllabic in final position.

where W is a weak cluster and $\hat{\Sigma}$ a stressed syllable, that is, a string of the form $C_0\acute{V}C_0$. Making minimal assumptions about surface structure, as before, this provides derivations such as (46):

(46)

$[_N photo\ [_{STEM} graph]_{STEM}\ ]_N$

1 — RULE (39), CASE (40d)

1 2 — RULE (45)

1 3 — RULE (43)

Where the stem (or inner noun) is polysyllabic, the stressed syllable will not be final and rule (45) will not apply. This accounts for the fact that in the examples in the right-hand column of (44), primary stress remains on the stem (or inner noun).[25]

Before proceeding to investigate other applications of rule (45), we can observe that it obviously falls together with rule (39). Combining (39) and (45), then, we have the following rule:

(47) $$V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0(W)\ /\ \text{——} (\begin{pmatrix} +\text{affix} \\ \hat{\Sigma} \end{pmatrix})]$$

where W is a weak cluster, C_0 is a string of zero or more consonants, $\hat{\Sigma}$ is a syllable of the form $C_0\acute{V}C_0$, and X does not contain # boundary internally. We will henceforth refer to this rule, with its various elaborations, as the Main Stress Rule, since it is the main rule applying to lexical categories. We return to this matter in Chapter Three.

In accordance with our notational conventions, rule (47) is an abbreviation for the sequence of rules:

(48)

(a) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0W\text{+affix}]$
(b) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0\text{+affix}]$
(c) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0W\hat{\Sigma}]$
(d) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0\hat{\Sigma}]$
(e) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0W]$
(f) $V \rightarrow [1\ \text{stress}] \ /\ X \text{——} C_0]$

Cases (a), (b), (e), (f) are, respectively, cases (a)–(d) of (40). As before, they constitute a disjunctively ordered block; if one of the four cases of (40) applies, none of the later ones is applicable. Furthermore, the notational conventions that we have given imply that if case (48c) applies, then case (d) is inapplicable, and that if either case (c) or (d) applies, then cases (e) and (f) are inapplicable. There are no further disjunctive constraints. The only permitted sequences of applicable rules, then, are the following:

(49)

(a), (c)
(a), (d)
(b), (c)
(b), (d)

Apart from these possibilities, at most one of the rules of (48) can apply. The order in which they become applicable is, aside from this restriction, the linear order of (48). These empirical assumptions follow from the general hypothesis regarding notations and the fact that (47) is the optimal representation of the processes so far discussed (see note 20).

[25] We have not yet given the rules that assign primary stress to these stems and inner nouns in the first cycle.

Before we continue with the analysis of English stress placement, let us make quite clear the status and character of our assumptions concerning the organization of grammars and the conditions on the applicability of grammatical rules. We have, so far, placed the following conditions on the grammar. The grammar is a linear sequence of rules of the form illustrated in (48), applying in accordance with the principle of the transformational cycle (see (15)). The relation of disjunctive ordering is defined on certain pairs of rules of this sequence by virtue of their formal similarities. To determine disjunctive ordering, we apply to the fullest possible extent the notational conventions involving parenthesization, bracketing, and the slash-dash notation defined as in (30)–(34). In this way we form an underlying schema which represents this sequence of rules and which is expandable into this sequence by the successive application of conventions involving the notations. (When this process is formalized later in our discussion, we will guarantee that the order of expansion is unique.) If at some stage in the expansion we reach a schema of the form $Z(X)Y$, expandable into the sequence of schemata ZXY, ZY, then all rules derived by expanding ZXY (or ZXY itself, if it is a rule) are disjunctively ordered with respect to all rules derived by expanding ZY (or ZY itself, if it is a rule). In this way, disjunctive ordering is defined on the rules of the sequence constituting the grammar. Notice that rules may be disjunctively ordered with respect to one another even if they are not adjacent in the ordering; for example, in (48), rule (a) is disjunctively ordered with respect to rule (f), but not with respect to rule (c).

The conventions associated with disjunctive ordering make use of the notations for stating grammatical schemata in a way that is rather novel within the theory of generative grammar. In earlier work these notations have been regarded solely as part of the system for evaluating grammars. They have been proposed as an explication of the notion "linguistically significant generalization"; the degree of linguistically significant generalization attained by a grammar—its "simplicity," in a technical sense of the term—is measured by the number of symbols appearing in the underlying schema that expands to this grammar by the use of the notations. (See Chomsky (1965) and many earlier references for discussion.) But now we are also making use of the notations to determine how the rules apply, in particular, to determine disjunctive ordering. That is to say, we are proposing that certain formal relations among rules, statable in terms of the notations that are used for the evaluation of grammars, are significant in determining how the grammar generates derivations. If the empirical hypothesis embodied in the definition of "disjunctive ordering" is correct, then this fact offers a powerful argument in support of the empirical reality of the evaluation procedures that have been developed within the theory of generative grammar, as it has evolved in recent years.

We can now return to the role of the Stressed Syllable Rule, as we shall henceforth refer to it—namely, cases (c) and (d) of the Main Stress Rule. We will refer to cases (a) and (b) of (48) as the Affix Rule.

Consider now the following sets of words:

(50)

torment	*torment*	*torrent*
convict	*convict*	*verdict*
export	*export*	*effort*
progress	*progress*	*tigress*

The words in the left-hand column are verbs, with stress on the final syllable; those in the other two columns are nouns, with primary stress on the penultimate syllable. Comparing the words in the middle column with those in the right-hand column, we can see that they

differ in the degree of stress on the final syllable and, concomitantly, in the quality of the final vocalic nucleus, which is reduced to [ə] in the right-hand column but not in the middle column.

We can account for the nouns in the middle column, that is, those with stress contour 13, by regarding them as derived from the corresponding verbs. Thus we view the relation between *torment* and *torment* as roughly analogous to the relation between *advertisement* and *advertise* or *impression* and *impress*. We then have derivations such as the following:

(51)

$[_N\,[_V$*torment*$]_V\,]_N$	
1	RULE (47), CASE (48f)
1 2	RULE (47), CASE (48d)
1 3	RULE (43)

In the first cycle, the Main Stress Rule applies to the underlying verb, assigning primary stress in the final strong cluster. Since the verb undergoes no further applications of the Main Stress Rule, in isolation it retains primary stress in this position. But the derived noun must undergo a second application of the Main Stress Rule, in accordance with the principle of the transformational cycle. In this application, the Stressed Syllable Rule applies, shifting primary stress to the left. Secondary stress on the final syllable is then weakened to tertiary by the Stress Adjustment Rule, giving the contour 13. The distinction between the elements of the left and middle columns of (50) can thus be attributed to the extra cycle in the derivation of the nouns. The distinction between the elements of the middle and right columns can be attributed to the fact that the right-hand elements are not derived from associated verbs and therefore have never received primary stress on the final syllable.[26] In this way, the Stressed Syllable Rule accounts for a distinction between tertiary and zero stress in the final syllables of pairs such as *torment–torrent*, *export–effort*.[27]

We have not yet explained why stress falls on the final syllable of the verb *progréss* in (50), even though this contains a weak cluster. As we will show in Chapter Three, Section 10, we must assume there to be a special boundary in such verbs—between *pro* and *gress* in this case—which blocks the application of (48e) in the first cycle but not of (48d) in the second cycle. Thus the derivation of the noun *prógress* from the underlying verb *progréss* will be identical to that of *tórment* in (51).

We have now seen two rather different effects of the Stressed Syllable Rule. In the case of *photograph* versus *photosynthesis*, it accounts for the distinction between a falling

[26] We have not yet given the rule that determines stress placement in nouns such as those of the right-hand column of (50). The fact is that in nouns, as distinct from verbs and adjectives, a final syllable with a simple vocalic nucleus is disregarded for purposes of stress placement, and the Main Stress Rule is then applied to the residue in the usual way. Thus, for nouns, a final syllable with a simple vocalic nucleus is treated in the same way as an affix and a stressed syllable by rule (47). We do not give this rule here because it involves certain assumptions with respect to notations and ordering that we prefer, for expository reasons, to leave for the next chapter. The facts are clear, however. By extending the Main Stress Rule in this way, we can account for the fact that primary stress appears in the penultimate syllable in the nouns of the rightmost column of (50), as well as in words such as *phílOgiston* and *horízon*, which have a strong medial cluster; that it appears in the antepenultimate syllable in words such as *vénison*, *cánnibal*, *élephant*, with a weak medial cluster and simple vocalic nucleus in the final syllable; and that it falls on the final syllable (by rule (48f)) in words such as *machine*, *careèr*, which have a complex vocalic nucleus in the final syllable.

[27] Observe that in the case of *torrent*, we know that the vowel of the final syllable is *e* (cf. *torrential*). In the case of *effort* there is no way of determining the phonological quality of the underlying vowel, which need not, therefore, be specified in the lexical entry for this formative.

and a rising contour for the prefix-stem combination, exactly as in the case of the noun *export* versus the verb *export*; in the case of *export* versus *effort* or *torment* versus *torrent*, it accounts for the difference between tertiary and zero stress in the final syllables.

Consider now words such as:

(52) (a) *relaxátion, annexátion, emendátion, connectívity, domesticity, authenticity*
(b) *devastátion, demonstrátion, contemplátion, opportúnity*

Observe that in each case the cluster preceding the primary stress is of the form VC_2 and is therefore a strong cluster, and that in each case this syllable has a weak stress.[28] However, the vowel quality is retained in the syllable preceding primary stress in the examples of (52a) but is lost in the same position in the examples of (52b). This distinction is clearly traceable to the fact that the examples of (52a) are derived from underlying forms in which this vowel has primary stress, whereas the examples of (52b) are derived from underlying forms in which this vowel is unstressed. Thus we have derivations such as the following:[29]

(53) (a) $[_N\,[_V relax]_V\ At+ion]_N$

1	RULE (47), CASE (48f)
2 1	(SEE NOTE 29)
2 3 1	(SEE NOTE 29)
3 4 1	RULE (43)

(b) $[_N\,[_V devastAt]_V\ ion]_N$

1 2	(SEE NOTE 29)
2 1	(SEE NOTE 29)
3 1	RULE (43)

Although certain details are not given in these derivations, there is still sufficient information to account for vowel quality in the weak-stressed syllable preceding primary stress. It is clear that the process of vowel reduction depends in a fundamental way on stress; in particular, a vowel that is sufficiently stressed, in some sense that we will make precise later, is protected from vowel reduction. Thus the degree of stress on the final syllable of *torment* (see derivation (51)) is sufficient to prevent vowel reduction, but that on the final syllable of *torrent* is not. Similarly, the second syllable of *relaxation*, having received primary stress in the first cycle, is immune to vowel reduction, but the second syllable of *devastation*, never having received any stress, does undergo the process of vowel reduction. In this way, we can account quite readily for the distinction between the examples of (52a) and (52b).

For some dialects (in particular, our own), we can find near minimal pairs to illustrate these far-reaching phonetic effects of the rules of the transformational cycle. Consider,

[28] Here, as elsewhere, we rely on the phonetic representations in Kenyon and Knott, which agree with our own pronunciation, with the provisos stated elsewhere. The stress on the syllable preceding primary stress cannot be stronger than [4 stress] in any of these cases, since the first syllable in each case has tertiary stress and the second (pre-main-stress) syllable is clearly weaker than the first. We would give the contour 3415 for (52a) and 3515 for (52b).

[29] These derivations involve various principles that will not be discussed until the next chapter. In particular, the affix *-ion* invariably places stress on the syllable immediately preceding it, and there is a rule changing a "21 contour to 231, as a special case of more general processes that we will discuss. We also omit here the rules that assign the proper stress contour 1"2 (which would become 1"3 by the Stress Adjustment Rule) to *devastAt* in the first cycle. Filling in these omissions will lead to no change in the analysis of the facts under discussion here.

for example, the words *compensation–condensation*.[30] In *condensation*, the vowel in the second syllable has received stress in the first cycle of the derivation because of the underlying verb *condense*: therefore, it does not reduce, and we have the phonetic representation [kandensAšən]. The corresponding vowel of *compensation*, never having received stress, is subject to vowel reduction, resulting in the phonetic representation [kampənsAšən].

To conclude this preliminary discussion of the principles that determine stress contours and the related phenomenon of vowel reduction, let us turn to the set of words in English that have the noun-forming affix *-y* (not to be confused with the adjective-forming *-y* of such words as *stringy* and *brawny*, which has very different phonetic effects and a different underlying representation). This is the affix that we find in such words as *aristocrac+y*, *econom+y*, *galax+y*. Before turning to its effect on stress placement, let us consider its phonological representation.

Phonetically, this affix is either [i] or [E], depending on the dialect; that is, it is a high front vowel of dialectally varying degree of tenseness and diphthongization. The tenseness and diphthongization give no information about the underlying phonological representation since there are no relevant contrasts in this position. As we shall see in the next chapter, even phonologically nontense vowels (i.e., simple vocalic nuclei) become tense and diphthongized in final position in the dialects in question. But, in fact, we do know that phonologically the affix cannot consist of a complex vocalic nucleus [E] if it is to be subject to the Main Stress Rule (47), since the cases of this rule that involve affixes, as we shall see, are restricted to affixes with simple vocalic nuclei.

With this possibility eliminated, let us now ask whether the affix *-y* can be phonologically represented as the simple vocalic nucleus *i*. An argument against this analysis is provided by consideration of the stem-forming vowel [i], which, along with the parallel stem-forming vowel [u], appears in the derived forms of pairs such as *proverb–proverbial*, *professor–professorial*, *habit–habitual*, *tempest–tempestuous*. The underlying forms must be represented in the lexicon in such a way as to indicate that they take the stem-forming augment [i] or [u] in their derived forms. A natural, and apparently the simplest, proposal is to enter these words in the lexicon in the form *professor+i*, *habit+u*, etc., with the augment deleted in final position by rule (54):

(54) $\begin{pmatrix} i \\ u \end{pmatrix} \rightarrow \phi \ / \ + \text{——} \ \#$

But if this suggestion is followed, then words such as *economy* cannot be entered with the representation *econom+i* for the affix will be incorrectly deleted in final position by rule (54).

These considerations suggest that the representation of the affix *-y* in lexical entries should be *+y*. That is, it should be entered as a high front glide, which later becomes a

[30] The latter is the nominalized verb that means "act of condensing," not the noun that means "a condensed state or form" or "a condensed mass" and that, although in some way related to the verb *condense*, is not derived from it as is *condensation* in the first sense. Kenyon and Knott give only the form with unreduced second syllable for *condensation*, and give both reduced and unreduced variants for *compensation*, as well as for the underlying form *compensate*. There is well-known dialectal divergence in these positions. In general, with respect to phonetic minutiae of this sort, it is impossible to expect complete consistency between speakers or for one speaker at various times. Nor should it necessarily be assumed that the transcriptions suggested by phoneticians, at this level of detail, correspond in any very clear way to an acoustic reality. As pointed out in Section 2 of this chapter, we are concerned here with ideal forms that may undergo various modifications in performance and that may relate more closely to a perceptual than an acoustic reality.

vowel by an extremely simple rule. We shall see, in fact, that the required rule converting *y* to *i* falls together with other rules that are needed on independent grounds. Thus, in terms of its analysis into vowels and consonants, the word *economy* is of the phonological form VCVCVCC, consistent, in fact, with the orthographic representation.

Adopting this quite well-motivated proposal, let us now turn to the effect of the affix *-y* on stress placement. We have already provided one quite general rule describing the effect of an affix on the assignment of primary stress, namely, cases (48a) and (48b) of the Main Stress Rule (47). But the affix *-y* does not seem to fall under this generalization, as we can see by considering data of the sort presented in (55), where the symbols W, S, and A stand for syllables terminating in weak, strong, and arbitrary clusters, respectively, and where the formula to the left of the colon describes the underlying form of the examples to the right:

(55)
(a) ÁW+*y*: *economy, policy, aristocracy*
(b) #ÁS+*y*: *industry, galaxy, modesty*
(c) ÁWS+*y*: *orthodoxy, testimony, rhinoplasty, promissory, auditory*
(d) ÁŚS+*y*: *advisory, compulsory, refractory, trajectory*[31]

The examples of case (a) are in fact consistent with the assumption that *-y* is simply a regular affix subject to the Affix Rule that is part of the Main Stress Rule (47). Since the syllable preceding the affix contains a weak cluster, case (48a) of (47) will assign primary stress to the syllable preceding this cluster, in the usual way. The examples of (55b), however, appear to be inconsistent with this assumption. If *-y* were subject to the Affix Rule, then primary stress would be placed on the strong cluster immediately preceding the affix, in accordance with case (48b) of rule (47), whereas in these examples primary stress is actually on the syllable preceding this strong cluster. Examples such as these might lead one to suggest another rule, unique to the suffix *-y*, namely, the rule that this suffix places primary stress on the syllable preceding it by two. Under such a rule, the examples of (55a) and (55b) would be accounted for.

The forms in (55c), however, show at once that this new proposal is incorrect. In these examples, primary stress is three syllables removed from the affix *-y*, and there is an unexplained tertiary stress on the syllable immediately preceding this affix (a syllable which, we observe, contains a strong cluster). We cannot simply add a special case requiring that stress be three syllables removed when *-y* is preceded by a strong cluster, for this possibility is excluded by the examples of (55d).

With no further attempt at patchwork solutions, let us see how close we can come to the facts by making the weakest and most general assumption, namely, that *-y* is simply a regular affix obeying the Main Stress Rule as it now stands.

As we have already noted, the examples of (55a) are consistent with this analysis. That is, the affix *-y* will now, like all affixes, assign stress to the syllable preceding a final weak cluster.

Consider next the examples of (55b). Under the assumption that *-y* is a regular affix, case (48b) of the Main Stress Rule (47) will place primary stress on the final syllable of the string preceding *-y*, since this syllable contains a strong cluster. This gives, for example, the form *industry*. Recall that according to the ordering constraints on the subcases (48a–f)

[31] We assume here that these words have the same affix *-Or+y* as *promissory, auditory*. Other analyses might be suggested for many of these words, taken in isolation, but the analyses we are supposing are at least as well motivated, on grounds independent of stress placement, as any others. We shall see directly that considerations of stress placement strongly support the analyses proposed here.

of rule (47), after (48b) has applied, case (c) or (d) may still be applied (see (49)). Case (48d) applies to a string of the form $VC_0\acute{\Sigma}]$, where $\acute{\Sigma}$ is a stressed syllable, assigning primary stress to the vowel. But, as we have noted above, the affix *-y* is a glide in the underlying representation. Hence *industry* is a string of the form $VC_0\acute{V}C_0]$, which is a special case of $VC_0\acute{\Sigma}]$. Case (48d) thus applies to *industry*, giving the stress pattern *industry*, after which the Stress Adjustment Rule applies to give *industry*. Other rules, to which we return below, determine that a tertiary-stressed vowel in the context of the *u* of *industry* loses its stress and reduces. This gives the desired stress pattern. The examples of (55b), then, are quite consistent with the assumption that *-y* is a regular affix.

Consider now the forms of (55c), which, as we have noted, are inconsistent with the assumption that *-y* places primary stress two syllables back. Taking *orthodoxy* as a typical example, the Main Stress Rule, as it stands, provides the following derivation:

(56)

$[_N\,[_A ortho\ [_{STEM} dox]_{STEM}\,]_A\ y]_N$		
	1	RULE (47), CASE (48f)
1	2	RULE (47), CASE (48c)
2	1	RULE (47), CASE (48b)
1	2	RULE (47), CASE (48c)
1	3	RULE (43)

In the first cycle, primary stress is placed on the monosyllabic stem *dox* (exactly as it is placed on the monosyllabic stem *graph* in the derivation (46) of *photograph*). In the next cycle we consider the adjective *orthodox*. The Stressed Syllable Rule (48c) places primary stress on the syllable preceding the weak cluster, again exactly as in the case of *photograph*. Thus, in isolation, the adjective would have the stress contour *orthodox* (the Stress Adjustment Rule weakening the final stress to tertiary). But in (56) there is still another cycle. In this third cycle, primary stress is assigned by the Affix Rule (48b) to the syllable with the strong cluster preceding the affix. The result is a string terminating with the stressed syllable *doxy*, a syllable of the form CVCCC. Hence the Stressed Syllable Rule (48c) applies once again, as it did in the preceding cycle, reassigning primary stress to the first vowel. The Stress Adjustment Rule (43) now applies to give the desired form *orthodoxy*. The other examples of (55c) are similar. In sum, these forms are consistent with the assumption that *-y* is a regular affix. The examples in (55d) are derived in a manner parallel to that of (55b), with case (48d) of the Stressed Syllable Rule applying on the last pass through the transformational cycle.

We see, then, that by taking the affix *-y* to be nonvocalic phonologically, all of the cases of (55) are explained on the assumption that it is a perfectly regular and unexceptional affix subject to the general Main Stress Rule. This fact alone would motivate the representation of the affix *-y* as a glide in underlying forms, but, as we have seen, there is independent support for this conclusion. The peculiar arrangement of data noted in (55) follows from this assumption, with no modification of the general rules. Here, then, is a striking example of the effectiveness of the principle of the transformational cycle, in conjunction with the principle of disjunctive ordering, in explaining otherwise quite refractory data.

Other forms in *-y* support these conclusions. Before turning to them, however, let us consider the following:

(57) *invéstigative, génerative, illústrative, demónstrative*

Clearly these have the underlying forms:

(58) *invéstigAt, génerAt, íllustrAt, démonstrAt*

But notice that the affix *-ive* should assign primary stress to the final strong syllable *-At*, in each case, giving the incorrect forms **investigátive*, **generátive*, **illustrátive*, **demonstrátive*. What actually happens is that the affix *-ive* assigns primary stress to the syllable immediately preceding *-At* if that syllable has a strong cluster, or one syllable further back if the syllable preceding *-At* has a weak cluster. In other words, primary stress is assigned just as if the affix were not *-ive*, but rather *-Ative*. In fact, we shall see that in general the element *-At* is considered to be a part of the affix for the purposes of stress placement. We can achieve this effect by reformulating the Main Stress Rule (47) as:

(59) $V \rightarrow [1\ \text{stress}] \ / \ X \text{——} C_0(W) \ / \ \text{——}((At)\begin{pmatrix}+\text{affix}\\ \acute{\Sigma}\end{pmatrix})]$[32]

To resolve an ambiguity in the expansion of the schema (59), let us assume, as a general principle, that braces are expanded before parentheses. With this assumption, schema (59) expands to (60), which is then expanded to a sequence of rules in the usual way.

(60)
$$V \rightarrow [1\ \text{stress}] \ / \ X \text{——} C_0(W) \ / \ \text{——}\begin{pmatrix}At+\text{affix}] \\ +\text{affix}] \\ At\ \acute{\Sigma}] \\ \acute{\Sigma}] \\ ]\end{pmatrix}\begin{matrix}\text{(a)}\\ \text{(b)}\\ \text{(c)}\\ \text{(d)}\\ \text{(e)}\end{matrix}$$

Disjunctive ordering holds between (60a) and (60b), between (60c) and (60d), and between each of (60a)–(60d) and (60e).

Let us consider the effect of this slight modification of the rule on examples with the affix *-y*. We will now have typical derivations such as (61) and (62), for *confiscatory* (similarly, *compénsatory*, *reformatory*, etc.) and *anticipatory* (similarly, *revérberatory*, *conciliatory*, etc.), respectively:

(61) $[_A\ [_V confiscAt]_V\ Or+y]_A$

1	2		(RULES TO BE GIVEN)
2	3	1	AFFIX RULE (60b)
3	1 4	2	STRESSED SYLLABLE RULE (60c)
4	1 5	3	RULE (43)

(62) $[_A\ [_V anticipAt]_V\ Or+y]_A$

1	2		(RULES TO BE GIVEN)
2	3	1	AFFIX RULE (60b)
1	4	2	STRESSED SYLLABLE RULE (60c)
1	5	3	RULE (43)

The two derivations correspond point by point. In both cases the stress contour is assigned to the underlying verb by rules that we will give later on. The verbs, in isolation, would be *confiscate*, *anticipate*. In the second cycle, the Affix Rule shifts primary stress to the strong

[32] Notice that the ordering implied by the use of parentheses carries over to this case, as we would expect. Thus, if the Affix Rule applies in the context ——*At*+*ive* (giving, e.g., *illustrAtive*), it is not permitted to reapply in the context ——*ive* (giving **illustrAtive*).

syllable immediately preceding the affix *-y*, in the usual way. At this point the Stressed Syllable Rule applies, under the modification (59)-(60)—that is, with the element *-At* regarded as part of the context of application rather than as subject to the application of the rule. Excluding *-AtOry* from consideration in this way, the rule assigns primary stress to the final strong syllable of the residual string *confisc-* in (61), and to the syllable preceding the final weak syllable of the residual string *anticip-* in (62). Stress is then weakened and vowels reduced in accordance with fairly straightforward rules to which we will return. Here, again, the various cases of the Main Stress Rule interact to generate some rather complex phonetic structures, in accordance with the general principle of the transformational cycle and the general empirical assumptions regarding ordering that we have formulated.

6. Particular and universal grammar

In Section 2, on the basis of some preliminary observations about stress contours in English, we suggested that certain principles of organization of a grammar might serve as preconditions for language acquisition, and we discussed some questions of psychological and physical fact relating to this assumption. Now, after a more detailed account of English stress contours, the tentative conclusions of Section 2 have been strengthened.

We have seen that simple rules applying under very general conditions can explain data of a rich and varied sort. This fact raises interesting and important questions. To facilitate the discussion of these questions, we can invoke a traditional distinction between "particular grammar" and "universal grammar." A particular grammar for a single language is a compendium of specific and accidental (that is, nonessential) properties of this language. A universal grammar is a system of conditions that characterize any human language, a theory of essential properties of human language. It is reasonable to suppose that the principle of the transformational cycle and the principles of organization of grammar that we have formulated in terms of certain notational conventions are, if correct, a part of universal grammar rather than of the particular grammar of English. Specifically, it is difficult to imagine how such principles could be "learned" or "invented" in some way by each speaker of the language, on the basis of the data available to him.[33] It therefore seems necessary to assume that these principles constitute a part of the schema that serves as a precondition for language acquisition and that determines the general character of what is acquired. While the general principles of organization of a grammar that we have been discussing can most plausibly be regarded as part of universal grammar, it seems that such rules as the Main Stress Rule must, in large part at least, be a part of the particular grammar of English. A reasonable tentative assumption, then, is that the Nuclear Stress Rule, the Compound Rule, and the Main Stress Rule must be learned by the child acquiring the language, whereas the conditions on the form of rules, the principle of the transformational cycle, and the principles of organization embodied in the various notational conventions that we have established are simply a part of the conceptual apparatus that he applies to the data.

[33] Furthermore, insofar as phonetic transcription corresponds to a perceptual rather than an acoustic reality —see Section 2—departures from the rules are undetectable. Quite apart from this, it is difficult to imagine that adults, whose perceptual set is extremely strong and whose phonetic acuity is very limited, could note and correct deviations in low-level phonetic forms even where these do have a direct counterpart in the physical shape of the utterance.

The Nuclear Stress Rule, the Compound Rule, and the Main Stress Rule, in its various cases, assign primary stress in certain positions. A very small body of data concerning the position of primary stress in simple utterances is sufficient to justify these rules. Correspondingly, a small body of data of this sort might be sufficient to enable the language learner to postulate that these rules form part of the grammar of the language to which he is exposed. Having accepted these rules, the language learner can now apply the general principles of universal grammar to determine their effects in a wide variety of cases. As we have seen, very simple rules can have extremely complex effects when applied in accordance with these general principles. The effects in themselves might well be undetectable by the native speaker or the language learner. When they are determined by a framework of internalized general principles, they become quite accessible to him.

Phonetically untrained speakers of a language seem to find it quite easy to determine the position of main stress in simple utterances, but extremely difficult to trace complex stress contours in a detailed and consistent way. There is, furthermore, some doubt as to the physical reality of these contours, although there is no doubt that with phonetic training, a speaker of the language can identify stress contours and other phonetic details with reasonable consistency. These observations are just what we would expect, given the assumptions to which we have tentatively been led about universal and particular grammar. A small body of data relating to the position of main stress can lead to the formulation of the major stress placement rules. Their effects in complex utterances are determined by the universal unlearned principles of organization of a grammar. There is no need for the speaker or hearer to attend to these automatically determined aspects of an utterance, even where they are physically real; but with training, they can be brought to the level of awareness, whether or not they have acoustic reality. In particular, stress contours can be "heard" with a fair degree of consistency even though they may not correspond in detail to any physical property of utterances.

7. On the abstractness of lexical representation

The syntactic component of the grammar contains a lexicon which lists lexical items with their inherent properties, in particular, those phonological properties that are not determined by general rule. The considerations of the preceding sections suggest that these underlying forms will in general contain no indication of the stress contour of the items or of the distinction between reduced and unreduced vowels. In these respects the lexical representation of an underlying form will be very different from the phonetic representations of its variants in particular contexts. As we investigate further, we will find many more dramatic examples of this discrepancy between underlying forms and their phonetic realizations.

In note 26, we pointed out that the placement of primary stress in nouns is governed by the following rule (where V, is a simple vocalic nucleus):

$$(63) \qquad V \rightarrow [1 \text{ stress}] \ / \ X\text{——}C_0(W) \ / \ \text{——}V_sC_0]_N$$

This rule clearly falls together with the general Main Stress Rule, in a way which we will examine in the next chapter. As pointed out in note 26, it accounts for the stress placement in words such as *vénison*, *horÍzon*, *élephant*. To assign primary stress in these words, we

disregard the final simple vocalic nucleus with the consonants following it, and assign primary stress to the penultimate syllable of the residue if its final cluster is weak or to this final cluster itself if it is strong. Thus the rule is of precisely the sort with which we are now familiar. If the final syllable of a noun contains a complex vocalic nucleus, then rule (63) is inapplicable, and case (48f) of the Main Stress Rule applies in the usual way, placing primary stress in the final syllable of such words as *machíne*, *careér*.

Superficially, words ending in vowels seem to contradict this rule. Thus, in words such as *country*, *ménu*, *window*, the final vocalic nucleus is complex (namely, *E*, *U*, *O*, respectively) in many dialects. Nevertheless, it does not receive stress. This seems difficult to explain within our present framework until we observe that there is no contrast between simple and complex vocalic nuclei in word-final position (see p. 39). Consequently, there is no barrier to representing words such as *country*, *menu*, *window* in the lexicon with simple vocalic nuclei in final position. This will then make the forms subject to rule (63), which excludes the final syllable from consideration and then assigns primary stress to the residue in the usual way. A later rule will then determine the quality of the word-final vocalic nucleus. This later rule is well motivated, apart from any question of stress placement. Hence these words do not contradict rule (63).

Further investigation of final unstressed vowels reveals that there is a peculiar gap in the pattern. We do not at this point in the exposition have the means to justify this remark, but we will be able to show that of the six simple vocalic nuclei that might appear in final position, only *i*, æ, *u*, *o*, and ɔ do in fact appear. There are no examples with *e* as the final vowel of the lexical representation.

With these observations as background, let us return to the problem of stress placement. Consider the words *ellípse*, *eclípse*. If the lexical representation were *elips*, *eklips*, then rule (63) would apply, eliminating the final syllable from consideration (since it contains a simple vocalic nucleus) and assigning primary stress to the first syllable, giving **Élips*, **Éklips* as the phonetic forms. Recall, now, the remarks of the preceding paragraph. Suppose that we were to assign to these words the lexical representations *elipse*, *eklipse*, respectively. Rule (63) will exclude the final simple vocalic nucleus *e* from consideration and will assign primary stress to the strong cluster that precedes it, giving *elípse*, *eklípse*. To obtain the correct phonetic forms, we now add the *e*-Elision Rule (64) to the grammar:

$$(64) \qquad e \rightarrow \phi \ / \ \text{——}\#$$

This rule gives the correct final forms. It also explains the gap noted in the preceding paragraph. We see now that this gap is not in the underlying lexical representations but only in the phonetic output.

Rule (64), as we shall see, has independent motivation apart from the considerations just mentioned. As one further example, consider the word *Neptune* with the phonetic representation [néptUn].[34] The final cluster of the phonetic representation is strong and hence should receive primary stress by the Main Stress Rule. We cannot simply add a final *e* in the lexical representation here, as we did in the preceding examples, for if we were to enter *Neptune* in the lexicon as *neptUne*, primary stress would still be placed on the second syllable, this time by rule (63). The only apparent alternative is to enter *Neptune* with the lexical representation *neptune*, that is, with the simple vocalic nucleus *u* in the second syllable. Rule (63) will now assign primary stress in the first syllable since the

[34] We overlook dialectal variants for the time being.

second syllable contains a weak cluster. We now add the rule (65) (where C is a single consonant):

(65) $$u \rightarrow U \;/\; \text{——}CV$$

We thus have the following derivation:

(66)

neptune	
I	RULE (63)
U	RULE (65)
ϕ	RULE (64)

The final phonetic form is [néptUn], as required.

Rule (65) is, in fact, justified on independent grounds. Thus we find only phonetic [U], and not the other phonetic reflexes of underlying *u*,[35] in the context ——CV (e.g., *music, mutiny, mural*).

Here, as in the forms discussed previously in this section, we are again led to an underlying representation which is quite abstract (and which, once again, corresponds directly to conventional orthography).

Consider next verbs such as *caréss* and *haráss*.[36] The final syllable of the phonetic representations for these forms has a stressed weak cluster, which is contrary to what is asserted by the Main Stress Rule (47). Suppose, however, that we were to provide these words with the lexical representations *kVress, hVræss*, with V here standing for an unspecified simple vocalic nucleus.[37] The two final consonants now make the final cluster strong, and case (48f) of the Main Stress Rule will apply to assign primary stress on this final strong cluster. To obtain the correct forms, we need another rule, which we shall call the Cluster Simplification Rule, to delete one of the *s*'s:

(67) The first of two identical consonants is deleted.[38]

This gives us [kərés], [hərǽs] as the phonetic forms, eliminating another apparent exception to the stress placement rules.

Once again, we find that the rule that we postulate (in this case, rule (67)) is well motivated on independent grounds, as we see from considerations such as the following. Consider first words such as *cunning, currency*, and *mussel*, in which the phonetic reflex of underlying *u* in the first syllable is [ʌ] rather than [U] (see note 35). According to rule (65), underlying *u* should give phonetic [U] in the context ——CV, as in *punitive, mural, music*, and so on. We can prevent the application of this rule to forms like *cunning* by assuming double consonants in the underlying representations. These will then simplify by rule (67). Alternatively, we would have to assume a contrast between *u* and *U* in underlying representations. This is highly implausible, not only because of the examples already noted that motivate rule (65), but also because of the system of vowel alternations that we shall describe.

Observe next that in the near pair *music–mussel*, noted above, the form with phonetic [U] has a voiced medial consonant, whereas the form with phonetic [ʌ] has an unvoiced

[35] The simple vocalic nucleus *u* of underlying lexical representations generally becomes phonetic [ʌ] before consonants by general rules that we will describe later.

[36] The latter, with the phonetic representation [hərǽs]. An alternative form, [hǽras], will derive from the lexical representation *hærVs*.

[37] We return later to the precise content of this remark.

[38] Notice that this rule is not, strictly speaking, formulable within the framework that we have established up to this point. We will return to this matter.

medial consonant. Thus the contrast is between [Uz] and [ʌs] in intervocalic position. This correlation is general. We can account for it by postulating a rule that voices [s] medially, this rule applying prior to (67):

(68) $$s \rightarrow [+\text{voice}] \;/\; V\text{——}V$$

Given the rule (68), which we will make more exact later on, we have the derivations (69):

(69)

musik	*mussel*	
U		RULE (65)
	ʌ	(SEE NOTE 35)
z		RULE (68)
	ϕ	RULE (67)

The rule (68) is independently motivated by many considerations. Compare, for example, pairs such as *resent–consent, resist–consist*, in which the initial consonant of each of the stems *-sent* and *-sist* voices intervocalically but not postconsonantally. Such examples give even more direct justification for rule (67)—the rule deleting the first of two identical consonants. Thus consider words such as *dissemble, dissent*, with the prefix *dis-* (cf. *distrust, disturb*, etc.) and a stem beginning with *s*. Evidently, rule (67) is required to account for the fact that the medial cluster is phonetically a single consonant [s]; it is protected from voicing by (68) because of the final *s* of the prefix, in contrast with *resemble, resent*, etc. Similarly, we must rely on rule (67) to account for the fact that the prefix *ex-* is phonetically [ek] when the stem begins with an [s], as in *exceed* versus *extend*. Thus, several considerations converge to support the analysis proposed.

Consider next words such as *radium, medial* versus *radical, medical*. These examples have the complex nuclei [A], [E] in the context ——*C*i*V*, and the simple nuclei [æ], [e] in the context ——*CiC*. A great many examples of this sort, which we shall study in detail below, lead us to postulate rules which have the following effect (where C is a single consonant):

(70) $$\begin{Bmatrix} æ \rightarrow A \\ e \rightarrow E \end{Bmatrix} \;/\; \text{——}CiV$$

Notice that where the vowel in question is followed by a double consonant (*calcium, compendium*), it is not subject to rule (70) and therefore remains simple.

We now proceed to words such as *potassium, gymnasium, magnesium*. As in the case of *music–mussel*, we find that where we have unvoiced [s], here in the context ——*i*V, the vocalic nucleus preceding it is simple, but where we have voiced [z], the vocalic nucleus preceding it is complex. We can now account for this arrangement of data with underlying forms and derivations much like the following:

(71)

potæssium	*gimnæsium*	
	A	RULE (70)
	z	RULE (68)
ϕ		RULE (67)

Once again, we rely on rule (67), among others, in accounting for the relevant data.

Finally, notice that words such as *confetti, Mississippi, Kentucky* appear to violate rule (63), which assigns stress in the antepenultimate syllable of a noun that ends in a simple vocalic nucleus preceded by a weak cluster. We can avoid this violation of the rule by giving the lexical representations *kVnfetti, mississippi, kVntukki*, respectively. The penultimate syllable, being strong, will now take primary stress by rule (63). The double consonants

prevent the voicing of [s] by rule (68) and the change of *u* to [U] by rule (65). Rule (67) then simplifies them, as before. In further support of this analysis, we observe that, quite generally, medial obstruent clusters are unvoiced in English; correspondingly, in the positions where a double consonant must be postulated to account for peculiarities in stress placement, consonant quality, and vowel quality, as in the examples of this paragraph, it is with rare exceptions an unvoiced obstruent that appears.

To recapitulate, the *e*-Elision Rule (64), the Cluster Simplification Rule (67), and the others that we have discussed here form a mutually supporting system of rules that can be justified in a variety of independent ways and that account for a fairly extensive array of data. These rules lead us to postulate underlying forms which are quite abstract. Furthermore, these abstract underlying representations are, in general, very close to conventional orthography.

We will conclude with two more examples. Consider the word *giraffe*, phonetically [jəræf]. Here we have a stress on the final weak cluster. We can explain this by postulating the underlying lexical representation *giræffe*. The rule (63) of stress placement assigns primary stress to the penultimate syllable. By *e*-Elision and Cluster Simplification (note again that an unvoiced cluster is involved) we derive [giræf]. Clearly we must have a rule that softens *g* to [j] (and *k* to [š]) before nonlow front vowels, with qualifications to be added later.

(72) $\begin{Bmatrix} g \\ k \end{Bmatrix} \rightarrow \begin{Bmatrix} j \\ s \end{Bmatrix} \; / \; \text{——} \begin{Bmatrix} i \\ e \end{Bmatrix}$

With rule (72) and the general rule of Vowel Reduction, we derive [jəræf], as required. Alternatively, we might take the underlying representation to be *jVræffe*; there are other possibilities for deriving the phonetic form by regular processes.

Finally, consider the words *courage* [kʌrəj] and *courageous* [kərʌjəs]. Superficially, these seem to contradict the rules of stress placement and vowel quality that we have presented in this chapter. Suppose, however, that we were to take the underlying form to be *koræge*.[39] On this assumption, we have the following derivations:

(73)

koræge	*koræge+əs*	
I		RULE (63)
	I	RULE (47), CASE (48a)
	A	RULE (70)[40]
J	J	RULE (72)
ϕ	ϕ	RULE (64)[41]
ʌ		(SEE NOTE 39)
ə	ə ə	(VOWEL REDUCTION)

[39] In our discussion of the Rounding Adjustment Rule in Chapter Four, we shall show that lax back vowels become unrounded under certain conditions. A consequence of this rule is the shift *o* → ʌ, where [ʌ] is regarded as a lax unrounded back mid vowel, differing from [o] only in not being rounded.

Incidentally, a better representation would be *koræge*, where *e* stands for a symbol identical in its feature composition to *k* except that it appears in a lexically designated class of forms that undergo certain syntactic and phonological processes (i.e., they take derivational affixes of the Romance and Greek systems and undergo rules such as (72)). We return to this matter at the end of Chapter Four.

[40] Actually, we generalize (70) so that it applies in the context —— C$_\alpha$V, where α is a nonlow front vowel or glide, that is, [i], [e], [ɪ], [ɛ], or the corresponding glides. This is a simplification of the rule, in our terms, as we shall see.

[41] We generalize rule (64) so that it elides final *e* not only before word boundary, but also before any formative boundary. This, too, is a simplification in our terms, as we shall see.

In the case of *courage*, in isolation, primary stress is placed by the Noun Rule (63); in the case of *courageous*, by the Affix Rule (47), in the familiar way. The second syllable of *courageous* becomes a complex nucleus by rule (70), before the nonlow front vowel followed by another vowel. The consonant *g* then softens to [j] by rule (72), and the final *e* is elided. Vowel Reduction then gives the desired forms. Once again, a quite abstract underlying form, very similar to conventional orthography, accounts for the variant forms by rules of great generality and wide applicability.

There is, incidentally, nothing particularly surprising about the fact that conventional orthography is, as these examples suggest, a near optimal system for the lexical representation of English words. The fundamental principle of orthography is that phonetic variation is not indicated where it is predictable by general rule. Thus, stress placement and regular vowel or consonant alternations are generally not reflected. Orthography is a system designed for readers who know the language, who understand sentences and therefore know the surface structure of sentences. Such readers can produce the correct phonetic forms, given the orthographic representation and the surface structure, by means of the rules that they employ in producing and interpreting speech. It would be quite pointless for the orthography to indicate these predictable variants. Except for unpredictable variants (e.g., *man–men*, *buy–bought*), an optimal orthography would have one representation for each lexical entry. Up to ambiguity, then, such a system would maintain a close correspondence between semantic units and orthographic representations. A system of this sort is of little use for one who wishes to produce tolerable speech without knowing the language—for example, an actor reading lines in a language with which he is unfamiliar. For such purposes a phonetic alphabet, or the regularized phonetic representations called "phonemic" in modern linguistics, would be superior. This, however, is not the function of conventional orthographic systems. They are designed for the use of speakers of the language. It is therefore noteworthy, but not too surprising, that English orthography, despite its often cited inconsistencies, comes remarkably close to being an optimal orthographic system for English. Correspondingly, it would not be surprising to discover that an adequate theory of the production and perception of speech will find a place for a system of representation not unlike orthography, though there is, for the moment, little evidence that phonemic transcription is a "psychologically real" system in this sense.

It should also be observed that very different dialects may have the same or a very similar system of underlying representations. It is a widely confirmed empirical fact that underlying representations are fairly resistant to historical change, which tends, by and large, to involve late phonetic rules.[42] If this is true, then the same system of representation for underlying forms will be found over long stretches of space and time. Thus a conventional orthography may have a very long useful life, for a wide range of phonetically divergent dialects.

These observations suggest a description of the process of reading aloud that might, to first approximation, be described in the following way. We assume a reader who has internalized a grammar G of the language that he speaks natively. The reader is presented with a linear stretch *W* of written symbols, in a conventional orthography. He produces as an internal representation of this linear stretch *W* a string *S* of abstract symbols of the sort that we have been considering. Utilizing the syntactic and semantic information available to him, from a preliminary analysis of *S*, as well as much extra-linguistic information

[42] See Halle (1964), Kiparsky (1965), Postal (1968).

regarding the writer and the context, the reader understands the utterance, and, in particular, assigns to S a surface structure Σ.[43] With Σ available, he can then produce the phonetic representation of S and, finally, the physical signal corresponding to the visual input W. Clearly, reading will be facilitated to the extent that the orthography used for W corresponds to the underlying representations provided by the grammar G. To the extent that these correspond, the reader can rely on the familiar phonological processes to relate the visual input W to an acoustic signal. Thus one would expect that conventional orthography should, by and large, be superior to phonemic transcription, which is in general quite remote from underlying lexical or phonological representation and not related to it by any linguistically significant set of rules. On the other hand, for an actor reading lines in a language that he does not know, phonemic transcription should be much superior to conventional orthography, since it can be read without comprehension, whereas conventional orthography, being close to the linguistically significant system underlying ordinary speech, can be read only when the surface structure (including the internal structure of words) is known, that is, when the utterance is to some degree understood.

There are many interesting questions that can be raised about the development of systems of underlying representation during the period of language acquisition. It is possible that this might be fairly slow. There is, for example, some evidence that children tend to hear much more phonetically than adults. There is no reason to jump to the conclusion that this is simply a matter of training and experience; it may very well have a maturational basis. Furthermore, much of the evidence relevant to the construction of the underlying systems of representation may not be available in early stages of language acquisition. These are open questions, and it is pointless to speculate about them any further. They deserve careful empirical study, not only because of the fundamental importance of the question of "psychological reality" of linguistic constructs, but also for practical reasons; for example, with respect to the problem of the teaching of reading. These further topics, however, lie beyond the scope of this book.

8. Vowel alternations

We have already noted that simple and complex vocalic nuclei alternate in some way. Let us now consider these processes in more detail.

A comparison of words such as *profane–profanity, compare–comparative, grateful–gratitude, serene–serenity, appeal–appelative, plenum–plenitude, divine–divinity, derive–derivative, reconcile–conciliate*, and innumerable others suggests that the grammar must contain rules which have the following effect:

(74)
$$\begin{aligned} \mathbf{A} &\rightarrow \text{æ} \\ \mathbf{E} &\rightarrow \text{e} \\ \mathbf{I} &\rightarrow \text{i} \end{aligned}$$

The vowel in boldface stands for a complex vocalic nucleus in the first member of each pair, and for a simple vocalic nucleus in the second member of each pair. Furthermore, both the vowel quality and the stress placement in the first member of each pair seem to

[43] Obviously, it is an oversimplification to assume that conversion of W to S precedes the interpretive processes that assign Σ to S. There is no reason for this having to be the case, and such commonplace phenomena as proofreading errors suggest that in fact it is not the case.

require that the underlying form have the complex rather than the simple vocalic nucleus, that is, that the rule be (74) rather than (75):

(75)
$$\begin{aligned} \text{æ} &\rightarrow \mathbf{A} \\ \text{e} &\rightarrow \mathbf{E} \\ \text{i} &\rightarrow \mathbf{I} \end{aligned}$$

Thus we postulate underlying forms such as *profAn, serEn, divIn*,[44] which are stressed on the final complex nucleus by the Main Stress Rule (case (48f)). To account for the second members of the pairs, we apply rule (74) in the context (76) (where $\bar{\text{V}}$ stands for an unstressed vocalic nucleus):

(76)
$$\text{——}\,\text{C}_0\bar{\text{V}}\text{C}_0\text{V}$$

Superficially, the vowel alternations illustrated by (74) appear to be extremely complex and unsystematic. We have disguised this fact by our capitalization notation. Stated in terms of symbols that receive a direct phonetic interpretation, the rules in (74) appear as:

(77)
$$\begin{aligned} \bar{\text{e}}\text{y} &\rightarrow \text{æ} \\ \text{iy} &\rightarrow \text{e} \\ \bar{\text{a}}\text{y} &\rightarrow \text{i} \end{aligned}$$

where the symbols $\bar{e}$, $\bar{\imath}$, $\bar{a}$ stand for phonetically tense counterparts to e, i, æ.[45] These rules are extraordinarily complex in terms of the otherwise well-motivated feature system that we will develop below and in terms of any concept of complexity that seems to have any merit at all.

Compounding the problem is the fact that it is not enough to postulate the rules (74)-(77); it is also necessary to postulate the rules (75), which have precisely the opposite effect. To see this, consider words such as *various–variety, German–Germanic–Germanium, manager–managerial*. The underlying form for *vary* must be *vAri*, with a final simple vocalic nucleus. Stress placement will then be determined correctly by rule (48e). The final vowel is converted from *i* to [E] finally or before another vowel by the rule discussed on page 45 in connection with words such as *country, window*. But notice that under stress, in *variety*, the vowel in question becomes not [E] but [I]. Therefore we must have a rule converting *i* to [I] in this position. Consider next the triple *German–Germanic–Germanium*. The position of stress on the first member of this triple shows that the vocalic nucleus of its final syllable must be weak. The second member shows that it must be æ. The third member shows that this underlying æ becomes [A] by a rule of the form æ → *A* in certain contexts (see rule (70) and the discussion of *courage–courageous* on p. 48). Consider now *manager–*

[44] If we were restricted to lowercase Latin letters and to a principle of absolute linearity of spelling, we could not use this device and would have to find an alternative notation. The proper choice is obvious, in the light of the rules given above. We can represent *profAn, serEn, divIn* in the form *profæne, serene, divine*; stress placement will now be determined correctly by the Main Stress Rule (case (48e)); the simple vocalic nuclei will become complex in the context —— Ce by a rule rather like (65); and the final *e* will be elided by rule (64), giving the correct phonetic forms.

One might inquire whether this proposal is not after all correct, for the underlying representations. We have considered this possibility quite seriously, and it has something to recommend it. We reject it, however, in favor of the analyses with final complex nuclei in the underlying representations, for two reasons which will become clearer later on: first, the solution with final *e* is less highly valued in terms of the general measure of evaluation (complexity measure) that we will develop; second, we have not been able to find a simple system of rules that gives the required results in detail under this assumption.

[45] The phonetics is straightforward except with respect to postulation of the [æ]–[āy] relation, which begs a few questions to which we shall return in Chapter Four.

managerial. Considerations of stress and vowel quality show that the final vowel of *manager* must be a simple vocalic nucleus. This vowel becomes [E] in the context ——CiV; it must therefore be the vowel *e* (since æ becomes [A] and *i* becomes [I]). Many examples of this sort show that we must, in fact, set up rules with the effect of (75), in addition to rules with the effect of (74).

We have now reached a conclusion which is quite unacceptable. The rules (74) (= (77)) and (75) are extremely complex in themselves. It is evident, furthermore, that there must be some underlying generalization that accounts for the fact that the rules (74) and the rules (75) are precisely opposite in their effects. If we give the rules in the form (74), (75), there is no way to express this generalization. In brief, we have two extremely complex processes which are surely related, but related in some way which is not statable if these processes are described in the form (74), (75).

These considerations suggest very strongly that something is seriously amiss in the analysis we have been tacitly assuming, with the symbols A, E, I, O, U taken simply as informal notational abbreviations for complex nuclei of underlying forms.

Notice that the processes (74) and (75) involve alternations of two kinds, from a phonetic point of view. We can see this by considering the formulation (77) of (74). Clearly these rules affect both the complexity and the quality of the vocalic nuclei in question; that is, the complex nuclei become simple, and the vowel of the vocalic nucleus changes in quality as well. Let us consider these processes individually.

To begin with, let us disregard the question of vowel quality and consider the matter of complexity of the vocalic nucleus. We note at once that the presence of the *y*-glide correlates with tenseness of the vowel. We need therefore account only for the tenseness. The presence of the glide will then be determined by the Diphthongization Rule (78):

(78) $$\phi \rightarrow y \;/\; \breve{V} \text{——}$$

where $\phi \rightarrow y$ stands for "insert *y*" and where $\breve{V}$ is a tense vowel. (We shall see that this rule is, in fact, more general.) We may now assume that there are no postvocalic glides in underlying forms.

The examples that we have already given illustrate fairly adequately the general scope of the rules governing tenseness. Summarizing what we have observed, we can formulate the following rules, as a first approximation:

(79) $$V \rightarrow [-\text{tense}] \;/\; \text{——} C\breve{V}CV$$

(80) $$V \rightarrow [+\text{tense}] \;/\; \text{——} \begin{cases} \# & \text{(a)} \\ V & \text{(b)} \\ C \begin{Bmatrix} i \\ e \end{Bmatrix} V & \text{(c)} \end{cases}$$

The Laxing Rule (79) converts the tense vowels in the boldface positions of *gratitude* (cf. *grAtefal*), *serenity* (cf. *serEn*), *derivative*, (cf. *derIv*) to their lax counterparts. If the underlying forms are *grēt*, *serēn*, *derīv*, respectively, rule (79) will give the forms *græt(itude)*, *seren(ity)*, *deriv(ative)*, as required. On the other hand, the Tensing Rule (80) will apply in the following way: (a) in the context —— #, the final vowels of *country*, *window*, *vary*, etc., will become tense; (b) in the context ——V, the vowels in boldface in *various*, *variety*, *impious*, *piety*, etc., will become tense; (c) in the context ——CαV (where α is a nonlow nonback vowel), the vowels in boldface in *managerial*, *courageous*, *Canadian*, etc., become tense. In all three cases, the tense vowel is diphthongized by rule (78).

The rules (78)–(80), which are quite simple and straightforward, account for the complexity of the vocalic nuclei in all of the cases that we have considered. The problem of vowel quality still remains, however, for the tense vowels (the complex vocalic nuclei). At this stage of our analysis, the vowels in boldface in the words *grateful*, *serene*, *derive*, for example, will be [æy], [ēy], [iy], respectively, from underlying æ, *e*, *i*, by Tensing and Diphthongization. But the vocalic nuclei of these words should be [ēy], [iy], [āy], respectively. That is, we must add a Vowel Shift Rule which has the following effect on stressed vowels:

(81) $$\begin{aligned} \bar{æ} &\rightarrow \bar{e} \\ \bar{e} &\rightarrow \bar{i} \\ \bar{i} &\rightarrow \bar{a} \; (= \bar{æ}\text{—see note 45}) \end{aligned}$$

In other words, the rule (81) effects the shifts:

(82) $$\bar{æ} \rightarrow \bar{e} \rightarrow \bar{i} \rightarrow \bar{æ}$$

We shall see, in Chapter Four, that the Vowel Shift Rule can be stated in a very simple way, and, in fact, that it can be generalized beyond the class of examples that we have considered. With the Tensing and Laxing Rules, the Diphthongization Rule, and the Vowel Shift Rule, we have now fully accounted for the examples considered so far, as we can see by the following typical derivations:

(83)

profæn (*profane*)	
profǽn	MAIN STRESS RULE (48f)
profǽyn	DIPHTHONGIZATION (78)
proféyn	VOWEL SHIFT (81)

(84)

profænity (*profanity*)[46]	
profǽnity	MAIN STRESS RULE (48a)
profǽnity	LAXING RULE (79)

(85)

mænVger (*manager*)	
mǽnVger	MAIN STRESS RULE (63)
mǽnVjer	RULE (72)
mǽnəjər	VOWEL REDUCTION

(86)

mænVgeriæl (*managerial*)	
mænVgériæl	MAIN STRESS RULE (48a)
mænVjériæl	RULE (72)
mænVjériæl	TENSING RULE (80c,b)
mænVjéyriyæl	DIPHTHONGIZATION (78)
mænVjíyriyæl	VOWEL SHIFT (81)[47]
mænəjíyriyəl	VOWEL REDUCTION

[46] In these derivations, we omit all cycles except the last.

[47] Note that the Vowel Shift Rule is restricted to vowels that carry stress, though not necessarily primary stress.

The points to be noted are the following. Instead of the extremely complicated rules (74), (75), we now have the quite simple rules (78)–(81).[48] More important, we have succeeded in expressing the generalization underlying the rules (74) and their inverses, the rules (75). By extracting the Vowel Shift Rule from these processes, we are left with only rules (79) and (80) (the Tensing and Laxing Rules) as inverses. This is a bare and irreducible minimum. We have, in other words, avoided the absurdity of assuming that the processes stated as (74) and (75) have no relation to each other. We now have abstract underlying representations such as *profæn, serēn, derīv, mænVger*. Observe that the device of capitalization used earlier corresponds to the phonological category of tenseness at the level of lexical representation. Note also that in the case of an underlying tense vowel, the corresponding phonetic element will invariably differ from the underlying vowel either in quality (if it remains tense) or in tenseness. For example, corresponding to the tense vowel in the boldface position in the underlying representation *serēn*, we have either [īy] (in the word *serene*) or [e] (in the word *serenity*). Once again, the postulated underlying forms are systematically related to conventional orthography (see note 44) and are, as is well known, related to the underlying forms of a much earlier historical stage of the language. There has, in other words, been little change in lexical representation since Middle English, and, consequently, we would expect (though we have not verified this in any detail) that lexical representation would differ very little from dialect to dialect in Modern English. If this assumption proves to be correct, it will follow that conventional orthography is probably fairly close to optimal for all modern English dialects, as well as for the attested dialects of the past several hundred years.

Bringing this discussion to a close, we will show that entirely independent considerations also support the postulation of the Vowel Shift Rule (81) for modern spoken English. In Section 7 we discussed the Velar Softening Rule that converts *g* to [j] and *k* to [s] before nonlow front vowels, that is [i], [e], [ɪ], and [ɛ]. But consider words such as:

(87)
criticism–critical–criticize
medicine–medical–medicate

Using the symbol *c* to represent unvoiced velars in lexical entries that are subject to the Velar Softening Rule (72) (see note 39), we have the underlying representations *critic-*, *medic-* for the base forms of (87). Evidently the Velar Softening Rule must precede the Vowel Reduction Rule, since we have softening in the boldface position in *medicine* (before underlying *i*) but not *medical* (before underlying æ), although in both cases the vowel following the consonant in question is reduced to [ə] by Vowel Reduction. Under this assumption, the words *criticism* and *critical* also cause no difficulty. But consider the words *criticize* and *medicate*. In the case of *criticize*, we have velar softening before a vocalic nucleus which is phonetically [ɪ] (= [āy]); in the case of *medicate*, we do not have velar softening before a vocalic nucleus which is phonetically [A] (= [ēy]). In other words, we have softening before a low back vowel but not before a nonlow front vowel, which is precisely the opposite of what we would expect in terms of rules of otherwise great generality. The paradox is resolved, of course, by the Vowel Shift Rule. The underlying representation for *criticize* is *criticīz*, and the underlying representation for *medicate* is *medicæt* (as indicated in both cases by the spelling—see note 44). If Velar Softening applies not only prior to Vowel Reduction but also prior to Vowel Shift, then we will have softening in the case of *criticize*

[48] The sense in which the latter rules are much simpler will be explained later. We shall argue that this is the only sense of "simplicity" that is relevant to the choice of a grammar.

(before an underlying high front vowel) but not *medicate* (with an underlying low vowel after the *c*). After Velar Softening applies, the Diphthongization and Vowel Shift Rules convert *ī* to [āy] (giving [kritisāyz]) and æ to [ēy] (giving [medikēyt]); in our alternative notation, the Velar Softening, Diphthongization, and Vowel Shift Rules convert underlying *criticīz, medicæt* to phonetic [kritisIz], [medikAt], respectively.

There are many other examples of this sort, some of which we will discuss when we deal with vowel alternations more carefully in Chapter Four. For the present, we simply point out that these examples provide an independent justification for the Vowel Shift Rule, and show once again the necessity of postulating lexical representations of a quite abstract sort.