Chapter eight

PRINCIPLES OF PHONOLOGY

1. On the evaluation procedure and the form of phonological rules

In this chapter we survey the formal devices that we propose for phonological description. These formal devices are a part of the theory of language underlying the description of English that we have presented. There are several conditions of adequacy that they must meet, and there are functions of various sorts that they must fulfill. They must, for example, make it possible to present the data precisely and clearly. They must, moreover, permit us to formulate general statements about the language which are true and significant, and must provide a basis for distinguishing these from other generalizations which are false, or which are true but not significant. Thus, if our analysis is correct, the rules in Chapter Five represent true and significant generalizations; they characterize the native speaker's competence, his idealized ability to produce and understand an unlimited number of sentences. The theory of English of which this study presents a part is based on a certain set of data, but it goes beyond these data, as any grammar must, both in depth and in scope—in depth insofar as it expresses the facts that underlie the data, and in scope insofar as it deals with other potential data, with linguistic forms that we did not specifically consider, including indefinitely many that have never been produced.

In terms of the formal devices that we permit, there are many rules that can be formulated that are incorrect for English. Obviously, a proper choice of formal devices does not guarantee selection of the correct grammar. It is to be expected that there will be many grammars which are statable in terms of the given formal devices and which are all compatible with whatever data are available from a certain language; thus selection among the alternatives will require a procedure of evaluation of some sort. Certainly this is true of any proposal that can now be envisioned with regard to the formalism for the statement of grammars. Furthermore, with other formal devices than those we allow it is possible to express "generalizations" that are consistent with the data but that are not, we would maintain, linguistically significant. When we select a set of formal devices for the construction of grammars, we are, in fact, taking an important step toward a definition of the notion "linguistically significant generalization." Since this notion has real empirical content, our particular characterization of it may or may not be accurate as a proposed explication. This point is important but often overlooked, and it may be useful to touch on the matter briefly here.

To clarify the empirical status of the formal devices selected for the theory of language, it is helpful to set the problem within the framework of psychological theory. The child is presented with certain "primary linguistic data," data which are, in fact, highly restricted and degraded in quality. On the basis of these data, he constructs a grammar that

defines his language and determines the phonetic and semantic interpretation of an infinite number of sentences. This grammar constitutes his knowledge of his language. It will, in particular, specify that the primary linguistic data are, in large measure, ill-formed, inappropriate, and contrary to linguistic rule.

These rather obvious facts pose the problem to which the linguist addresses himself, that is, to account for the child's construction of a grammar and to determine what preconditions on the form of language make it possible. Our approach to this problem is two-pronged. First we develop a system of formal devices for expressing rules and a set of general conditions on how these rules are organized and how they apply. We postulate that only grammars meeting these conditions are "entertained as hypotheses" by the child who must acquire knowledge of a language. Secondly, we determine a procedure of evaluation that selects the highest valued of a set of hypotheses of the appropriate form, each of which meets a condition of compatibility with the primary linguistic data. We will not concern ourselves here with the nontrivial problem of what it means to say that a hypothesis—a proposed grammar—is compatible with the data, but will restrict ourselves to the other two problems, namely, the specification of formal devices and of an evaluation procedure. In other words, we make the simplifying and counter-to-fact assumption that all of the primary linguistic data must be accounted for by the grammar and that all must be accepted as "correct"; we do not here consider the question of deviation from grammaticalness, in its many diverse aspects. Given this simplifying assumption, we face the empirical problem of selecting a set of formal devices and an evaluation procedure which jointly meet the empirical condition that the highest valued grammar of the appropriate form is, in fact, the one selected by the child on the basis of primary linguistic data. Even with this idealization, a proposed theory that specifies formal devices and an evaluation procedure can be proven false (all too easily, in actual fact) by confronting it with empirical evidence relating to the grammar that actually underlies the speaker's performance. There is such a grammar, and it is an empirical problem to discover it and to determine the basis for its acquisition. However difficult it may be to find relevant evidence for or against a proposed theory, there can be no doubt whatsoever about the empirical nature of the problem. We stress this fact because the problem has so often been misconstrued as one of "taste" or "elegance."

A further word of caution is perhaps necessary in connection with this formulation of the general problems that guide our study of language. Apart from the idealization mentioned in the preceding paragraph, there is another, much more crucial, idealization implicit in this account. We have been describing acquisition of language as if it were an instantaneous process. Obviously, this is not true. A more realistic model of language acquisition would consider the order in which primary linguistic data are used by the child and the effects of preliminary "hypotheses" developed in the earlier stages of learning on the interpretation of new, often more complex, data. To us it appears that this more realistic study is much too complex to be undertaken in any meaningful way today and that it will be far more fruitful to investigate in detail, as a first approximation, the idealized model outlined earlier, leaving refinements to a time when this idealization is better understood. The correctness of this guess, of course, will have to be judged by the long-range effectiveness of a research program of this sort, as compared with alternatives that might be imagined. In the meantime, this idealization must be kept in mind when we think about the problem of the "psychological reality" of the postulated mental structures.

To take a concrete example, consider the matter of the synchronic residue of the

English Vowel Shift, discussed in detail in Chapter Six. We have argued that the underlying lexical forms in English contain vowels in pre-Vowel-Shift representation, and that these forms are what would have psychological reality given the other assumptions in our model—in particular, the assumption of instantaneous language acquisition. To the extent that these assumptions are false to fact, the conclusions that follow from them may also be false to fact. In particular, it is no doubt the case that the linguistic forms that justify our postulation of the Vowel Shift Rule in contemporary English are, in general, available to the child only at a fairly late stage in his language acquisition, since in large measure these belong to a more learned stratum of vocabulary. Since the order of presentation of linguistic data is, for the moment, an extrinsic factor that has no place in our theory, we cannot take account of this fact, and we can therefore state our conclusion about psychological reality only in hypothetical form: *if it were the case that language acquisition were instantaneous, then the underlying lexical forms with pre-Vowel-Shift representations would be psychologically real.* This, we propose, is a true statement about language—ultimately, about mental processes and the particular way in which they function. But an empirical conclusion of this sort will, naturally, be more difficult to verify, will require more indirect and subtle means of verification, than a simple categorical assertion. To us it seems that for the foreseeable future, the study of language and mental processes will have to be carried out at such a level of abstraction if it is to make significant progress.

With these background comments in mind, let us turn to the formal devices that we have been using in our exposition of English sound structure.

The rules that we assign to the phonological component have generally been presented in terms that can be symbolized by the formula:

$$(1) \qquad A \rightarrow B \;/\; X \text{——} Y$$

where A and B represent single units of the phonological system (or the null element); the arrow stands for "is actualized as"; the diagonal line means "in the context"; and X and Y represent, respectively, the left- and right-hand environments in which A appears. These environments may be null, or may consist of units or strings of units of various sorts, and may also include labeled brackets representing the syntactic category of the string to which the rule is applied.

Consider the hypothetical languages A and B which have identical phonological systems consisting of the vowels /i u æ a/ and the other phonological units shown in Table 1. Assume that language A has the rules of (2), whereas language B has the rules of (3).

TABLE 1. *The sound systems of languages A and B*[a]

	i	u	æ	a	r	l	p	t	k	s	m	n	y	w
vocalic	+	+	+	+	+	+	−	−	−	−	−	−	−	−
consonantal	−	−	−	−	+	+	+	+	+	+	+	+	−	−
high	+	+	−	−	(−)	(−)	(−)	(−)	(+)	(−)	(−)	(−)	(+)	(+)
back	−	+	−	+	(−)	(−)	(−)	(−)	(+)	(−)	(−)	(−)	−	+
anterior	(−)	(−)	(−)	(−)	−	+	+	+	−	+	+	+	(−)	(−)
coronal	(−)	(−)	(−)	(−)	(+)	(+)	−	+	(−)	+	−	+	(−)	(−)
continuant	(+)	(+)	(+)	(+)	(+)	(+)	(−)	−	(−)	+	(−)	(−)	(+)	(+)
nasal	(−)	(−)	(−)	(−)	(−)	(−)	−	−	(−)	(−)	+	+	(−)	(−)
strident	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(+)	(−)	(−)	(−)	(−)

[a] The meaning of the parenthesization will be discussed directly.

(2)

RULES OF LANGUAGE A

i → y / ——p
i → y / ——r
i → y / ——y
i → y / ——a

(3)

RULES OF LANGUAGE B

i → y / ——p
r → l / ——r
t → p / ——y
s → n / ——a

The difference between (2) and (3) lies in the fact that the statements in (2) are partially identical, whereas those in (3) are totally different from one another. This difference, which is clearly of linguistic interest, would be expressed if we introduced into our formalism a device akin to conjunction in ordinary English, which would permit us to coalesce two partially identical rules into a single rule without repeating the parts that are identical. We therefore establish the convention (4):

(4) Two partially identical rules may be coalesced into a single rule by enclosing corresponding nonidentical parts in braces: { }.

This convention enables us to rewrite (2) as (5):

(5)

$$i \rightarrow y \,/\, \text{——} \begin{Bmatrix} p \\ r \\ y \\ a \end{Bmatrix}$$

However, it does not permit (3) to be similarly abbreviated. Let us call (5) a "schema" which "expands" to the sequence of rules (2). The convention (4) is one of a set of notational conventions that allow certain sequences of rules (or schemata) to be abbreviated by schemata. In informal discussion, when no confusion can arise, we will not consistently maintain the distinction between the terms "rule" and "schema," extending "rule" to schemata as well.

Implicit in the brace notation is the assumption that languages tend to place partially identical rules such as those in (2) next to one another in the ordered sequence of rules that constitutes the phonological component of a grammar: it is only when partially identical rules are adjacent to one another that the brace notation can be exploited. It has been noted by Kiparsky (forthcoming) that phonological change provides evidence in support of this assumption. One of the examples discussed by Kiparsky is the evolution of the rules laxing vowels before consonant clusters (see (20III), Chapter Five) and in the pre-penultimate syllables of a word (see (20IV), Chapter Five). The historical antecedents of these two rules differed from their modern counterparts in that preconsonantal laxing took place before three (instead of two) or more consonants, whereas trisyllabic laxing required that the vowel be followed by two (instead of one) consonant. The historical change, then, was that both of these rules decreased by one the number of consonants that must follow the vowel to be laxed. This parallelism may be regarded as a mere coincidence, as has been the approach in every treatment of English historical phonology known to us. Alternatively, and

more satisfactorily, in view of the fact that there is no evidence to show that the changes in the two rules were due to separate processes, the parallelism may be regarded as being the result of a single change: the generalization of the schema (6) to the schema (7) by deletion of one of the consonants that must follow the vowel to be laxed.

(6) $$V \rightarrow [-\text{tense}] \; / \; \text{---} \; CC \begin{Bmatrix} C \\ VC_0V \end{Bmatrix}$$

(7) $$V \rightarrow [-\text{tense}] \; / \; \text{---} \; C \begin{Bmatrix} C \\ VC_0V \end{Bmatrix}$$

The characterization of the change as a single process, however, presupposes the existence of rule schemata as entities to which phonological changes may apply. Since schemata exist in a grammar only by virtue of conventions such as those discussed in this section, the examples just cited might be regarded as evidence in support of the reality of rule schemata and the conventions governing their use.

We can make use of notational conventions such as (4) to provide an evaluation procedure for grammars if we supplement the conventions with the following definition:

(8) The "value" of a sequence of rules is the reciprocal of the number of symbols in the minimal schema that expands to this sequence.[1]

where the minimal schema is the one with the smallest number of symbols. More generally, let us say that if the schema Σ_1 expands to the sequence of rules $R_1, \ldots, R_m$ and the schema Σ_2 expands to the sequence of rules $S_1, \ldots, S_n$, then the sequence of schemata Σ_1, Σ_2 expands to the sequence of rules $R_1, \ldots, R_m$, $S_1, \ldots, S_n$; and let us accept the analogous convention for a sequence of schemata $\Sigma_1, \ldots, \Sigma_p$ of arbitrary length. Let us now say that the "minimal representation" of a sequence of rules is the sequence of schemata with the smallest number of symbols that expands to this sequence of rules.[2] We can then restate definition (8) as (9):

(9) The "value" of a sequence of rules is the reciprocal of the number of symbols in its minimal representation.

Let us return now to the rules of (2) and (3). Given the conventions (4) and (9), the sequence of rules (2) is more highly valued than the sequence of rules (3): the minimal representation of (2) is (5) and the minimal representation of (3) is (3) itself, and (5) has fewer symbols than (3). Within the general framework of our theory, as described above, the conventions (4) and (9) imply that a linguistically significant generalization underlies (2) but not (3). Although in this case the fact may seem too trivial to require extensive comment, as we proceed further along the same lines we soon reach conclusions that are quite non-trivial, such as some of those discussed in the preceding chapters.

It should be observed in this connection that although definition (9) has commonly been referred to as the "simplicity" or "economy criterion," it has never been proposed or intended that the condition defines "simplicity" or "economy" in the very general (and

[1] We take the value to be $\frac{1}{n}$, where n is the number of symbols, so that the phrase "higher valued" will have its natural intuitive and numerical meaning.

[2] We give a more precise statement of these definitions in the Appendix to this chapter. Notice that the minimal representation may not be unique.

still very poorly understood) sense in which these terms usually appear in writings on the philosophy of science. The only claim that is being made here is the purely empirical one that under certain well-defined notational transformations, the number of symbols in a rule is inversely related to the degree of linguistically significant generalization achieved in the rule. In other words, definition (9), together with a specific choice of an alphabet from which the symbols are selected (see Section 2) and a specific set of notations for formulating rules and schemata, provides a precise explication for the notion "linguistically significant generalization" (Halle, 1962; Chomsky, 1964; Chomsky and Halle, 1965). Like all empirical claims, this can be tested for correctness and accuracy and can readily be controverted by evidence showing that it fails to hold true in certain clear cases.

2. *Segments as feature complexes*

We have as yet said nothing about the symbols that are used to represent the entities in our rules. In the present study speech sounds, or, more technically, segments, as well as all boundaries, are formally treated as complexes of features rather than as further unanalyzable entities. We assume, in other words, that the units or strings of units represented by the letters *A*, *B*, *X*, *Y* in (1) consist of feature columns or sequences of feature columns such as those shown in Table 1. The symbols referred to in the evaluation criterion (9) will, then, naturally be taken as distinctive feature specifications such as [+vocalic] or [−nasal].

The decision to regard speech sounds as feature complexes rather than as indivisible entities has been adopted explicitly or implicitly in almost all linguistic studies. Specifically, it is almost always taken for granted that phonological segments can be grouped into sets that differ as to their "naturalness." Thus, the sets comprising all vowels or all stops or all continuants are more natural than randomly chosen sets composed of the same number of segment types. No serious discussion of the phonology of a language has ever done without reference to classes such as vowels, stops, or voiceless continuants. On the other hand, any linguist would react with justified skepticism to a grammar that made repeated reference to a class composed of just the four segments [p r y a]. These judgments of "naturalness" are supported empirically by the observation that it is the "natural" classes that are relevant to the formulation of phonological processes in the most varied languages, though there is no logical necessity for this to be the case. In view of this, if a theory of language failed to provide a mechanism for making distinctions between more or less natural classes of segments, this failure would be sufficient reason for rejecting the theory as being incapable of attaining the level of explanatory adequacy.

Translated into feature terms, rule (5) will appear as (10) (p. 336), where the feature complexes representing a given unit (segment) are enclosed in square brackets: [ ].

By viewing segments as complexes of a fixed language-independent set of features, we have established a part of a mechanism which is required to distinguish more natural from less natural sets of segments; we can now say that sets of segments that have features in common are more natural than sets of segments that have no common features. What remains to be decided is the "metric of naturalness," that is, whether sets of segments sharing a large number of features are more "natural" than sets of segments sharing fewer features, or whether the reverse or perhaps some totally different relationship is the appropriate formal expression of this concept. Before making this decision, it is useful to state certain