1

From *The Sound Pattern of English*: Phonetic and Phonological Representation (1968)

Noam Chomsky and Morris Halle

. . . [T]he phonetic transcription is related by the rules of the phonological component to a string of formatives with labeled bracketing which represents the surface syntactic structure of the sentence. We will now examine in some detail the manner in which these formatives are represented in a linguistic description. Many of the formatives are lexical items, the "roots" or "stems" of traditional grammar. A grammar must include a list of these items, for part of a speaker's knowledge of his language consists of knowing the lexical items of the language. It is by virtue of this knowledge that the native speaker is able to distinguish an utterance in normal English from an utterance such as Carnap's "Pirots karulized elatically" or from Carroll's jabberwocky, which conform to all rules of English but are made up of items that happen not to be included in the lexicon of the language.

The representations of the individual items in the lexicon must incorporate the knowledge which makes it possible for the speaker to utilize each lexical item in grammatically correct sentences. This includes certain syntactic information which the speaker must have. For example, he must know that a particular item is a noun and that it belongs to a large number of intersecting categories such as "animate" or "inanimate," "human" or "nonhuman," "feminine" or "masculine." Since the only question of interest here is whether or not a given item belongs to the category in question, it is natural to represent this information by means of a binary notation: *cow*, for example, would be specified as [+animate, –human, +feminine]. In addition to these syntactic features, each lexical entry must contain specified features which determine the phonetic form of the item in all contexts. We shall call these the "phonological features." The phonological features cannot be chosen arbitrarily, for the phonological component would then have to include a huge number of ad hoc rules of the type

$$[+A, -B, -C, +D] \rightarrow [hʌ́t]$$
$$[-A, -B, -C, +D] \rightarrow [rʌ́t]$$
$$[-A, +B, -C, +D] \rightarrow [əlíps]$$

Moreover, if we represented lexical items by means of an arbitrary feature notation, we would be effectively prevented from expressing in the grammar the crucial fact that items which have similar phonetic shapes are subject to many of the same rules.

We might consider overcoming these difficulties by representing each lexical item in its phonetic representation. However, this solution is not open to us either, for a lexical item frequently has several phonetic shapes, depending on the context in which the item appears. If we chose to represent each lexical item by the set of its phonetic representations, we would be treating all phonetic variations as exceptions and would, in principle, be unable to express within our grammar the phonetic regularities and general phonological processes that determine phonetic form. If, on the other hand, we chose to allow only a single phonetic representation for each item, then we would have to provide some rationale for our selection. Furthermore, it is easily shown that many of the most general and deep-seated phonological processes cannot be formulated as rules that directly relate phonetic representations; rather, these processes presuppose underlying abstract forms.

We therefore can represent lexical items neither in phonetic transcription nor in an arbitrary notation totally unrelated to the elements of the phonetic transcription. What is needed is a representation that falls between these two extremes. Accordingly we propose that each item in the lexicon be represented as a two-dimensional matrix in which the columns stand for the successive units and the rows are labeled by the names of the individual phonetic features. We specifically allow the rules of the grammar to alter the matrix, by deleting or adding columns (units), by changing the specifications assigned to particular rows (features) in particular columns, or by interchanging the positions of columns. Consequently, the matrix that constitutes the phonetic transcription may differ quite radically from the representation that appeared in the lexicon. There is, however, a cost attached to such alterations, for they require the postulation of rules in the phonological component. Such rules are unnecessary in cases where the lexical representation can be accepted as the phonetic representation. In general, the more abstract the lexical representation, the greater will be the number and complexity of the phonological rules required to map it into a phonetic transcription. We therefore postulate abstract lexical entries only where this cost is more than compensated for by greater overall simplification – for example, in cases where the combination of abstract lexical entries and a set of rules permits the formulation of phonological processes of great generality that would otherwise be inexpressible.

Thus, lexical representations and a system of phonological rules are chosen in such a way as to maximize a certain property that we may call the "value" of the grammar, a property that is sometimes called "simplicity." As has been emphasized repeatedly in the literature, the concept of "simplicity" or "value" is an empirical one. There is some correct answer to the question of how lexical items are represented and what the phonological rules are. A particular notion of "value" or "simplicity" will lead to an assumption about lexical items and phonological rules which is either right or wrong, and therefore the validity of the notion must be determined on empirical grounds, exactly as in the case of every other concept of linguistic theory. It may be difficult to obtain crucial empirical evidence bearing on proposed definitions of "simplicity," but this cannot obscure the fact that it is an empirical concept that is involved, and that one can no more employ a priori arguments in determining how "value" should be defined than in determining how to define "set of distinctive features" or "grammatical transformation" or any other concept of linguistic theory.

A specific proposal as to the definition of "value" will make certain assumptions as to what constitutes a linguistically significant generalization, as to what constitutes a "regularity" of the sort that a child will use as a way of organizing the data he is confronted with in the course of language acquisition. The child is presented with certain data; he arrives at a specific grammar, with a specific representation of lexical items and a certain system of phonological rules. The relation between data and grammar is, we naturally assume, language-independent: there is no basis for supposing that individuals differ genetically in their ability to learn one rather than another natural language. Consequently, the relationship is determined by a principle of universal grammar. Specifically, the definition of "value" or "simplicity" must be part of universal grammar, and a specific proposal will be right or wrong as it does or does not play its part in accounting for the actually existing relation between data and grammar.

Summarizing, we postulate a set of lexical matrices and a system of phonological rules which jointly maximize value, in some sense which will be defined. Phonological representation in terms of lexical matrices (as modified through readjustment rules) is abstract in the sense that the phonological representation is not necessarily a submatrix of the phonetic representation. We do not, in other words, impose the conditions of linearity and invariance (see Chomsky, 1964) on the relation between phonological and phonetic representation. The indirectness of this relation must be purchased at the cost of adding rules to the grammar. Given a definition of "value," we can therefore say that the facts of pronunciation induce the representation of items in the lexicon.

Notice that the phonetic features appear in lexical entries as abstract classificatory markers with a status rather similar to that of the classificatory features that assign formatives to such categories as "noun," "verb," "transitive." Like the latter, the phonological features indicate whether or not a given lexical item belongs to a given category. In the case of the phonological matrices, these categories have the meaning "begins with a voiced stop," "contains a vowel," "ends with a strident nonback obstruent," and so on. In view of the fact that phonological features are classificatory devices, they are binary, as are all other classificatory features in the lexicon, for the natural way of indicating whether or not an item belongs to a particular category is by means of binary features. This does not mean that the phonetic features into which the phonological features are mapped must also be binary. In fact, the phonetic features are physical scales and may thus assume numerous coefficients, as determined by the rules of the phonological component. However, this fact clearly has no bearing on the binary structure of the phonological features, which, as noted, are abstract but not arbitrary categorial markers.[1]

As already noted, the phonetic representation can be thought of formally as a two-dimensional matrix in which the columns stand for consecutive units and the rows stand for individual phonetic features. The phonetic features can be characterized as physical scales describing independently controllable aspects of the speech event, such as vocalicness, nasality, voicing, glottalization. There are, therefore, as many phonetic features as there are aspects under partially independent control. It is in this sense that the totality of phonetic features can be said to represent the speech-producing capabilities of the human vocal apparatus. We shall say that the phonetic representations of

two units are distinct if they differ in the coefficient assigned to at least one feature; phonetic representations of sequences of units are distinct if they contain distinct units or if they differ in the number or order of units.

At the level of phonetic representation, utterances are comparable across languages; it thus makes sense to ask whether the phonetic representation of an utterance of language L_1 is distinct from a phonetic representation of an utterance of a different language L_2. For example, an utterance containing an apical dental stop must have a different phonetic representation from an utterance that is identical except for containing a laminal dental stop in place of the apical dental stop. The representation must differ, since the distinction is determined in part by language-specific rules; it is not a case of universal free variation. An interesting example of cross-language contrasts that require a special phonetic feature is provided by the labiovelar consonants found in many African languages. In some languages, such as Yoruba, these consonants are produced with a special clicklike suction, whereas in other languages, such as Late, they are produced without this suction (Ladefoged, 1964, p. 9). Since clicklike suction is clearly an independently controllable aspect of the speech event, the data just cited establish suction as a separate phonetic feature, regardless of the fact that apparently in no language are there contrasting pairs of utterances that differ solely in this feature.

The situation is not always straightforward, however. Since phonetic features are scales which may in principle assume numerous discrete coefficients, the question may arise, under certain circumstances, whether a certain phonetic contrast is to be represented by means of a new phonetic feature or by increasing the number of coefficients that some already extant phonetic feature may be allowed to assume. The latter solution may appear especially attractive in cases where a slight redefinition of some phonetic feature would readily accommodate the proposed solution.

To summarize, the features have a phonetic function and a classificatory function. In their phonetic function, they are scales that admit a fixed number of values, and they relate to independently controllable aspects of the speech event or independent elements of perceptual representation. In their classificatory function they admit only two coefficients, and they fall together with other categories that specify the idiosyncratic properties of lexical items. The only condition that we have so far imposed on the features in their lexical, classificatory function is that lexical representations be chosen in such a way as to maximize the "value" of the lexicon and grammar, where the notion "value" is still to be defined precisely, though its general properties are clear. Apart from this, the representation of a lexical item as a feature complex may be perfectly abstract.

In a later discussion (see chapter 9 [of *Sound Pattern of English*]), we will consider significantly heavier conditions on lexical representation. There we will turn to the question of "plausible phonological rules" and, more generally, to ways in which a particular feature may or may not function in the lexicon and in the phonology. These considerations will differentiate features from one another with respect to the role that they can play in the system of rules and in lexical representation. At that point in the development of our theory, considerations beyond maximization of value will enter into the determination of lexical representations.

Note

1 Failure to differentiate sharply between abstract phonological features and concrete phonetic scales has been one of the main reasons for the protracted and essentially fruitless debate concerning the binary character of the Jakobsonian distinctive features.

References

Chomsky, Noam. 1964. *Current Issues in Linguistic Theory*. The Hague: Mouton.

Ladefoged, Peter. 1964. *A Phonetic Study of West African Languages*. West African Language Monographs, 1. Cambridge: Cambridge University Press.