PHONOLOGICAL THEORY
Evolution and Current Practice

VALERIE BECKER MAKKAI
University of Illinois at Chicago Circle

HOLT, RINEHART AND WINSTON, INC.
New York  Chicago  San Francisco  Atlanta  Dallas
Montreal  Toronto  London  Sydney
THE ROLE OF
A PHONOLOGICAL FEATURE SYSTEM
IN A THEORY OF LANGUAGE*
JAMES D. MCCAWLEY

The system of phonological features used in Halle's *The Sound Pattern of Russian* is essentially the same as that presented in Jakobson, Fanig, and Halle's *Preliminaries to Speech Analysis* eight years earlier. It is important to note, however, that the system of features played a very different role in the theoretical frameworks of these two works. In *Preliminaries*, as in Jakobson's other phonological writings, the features are intended as distinctive features: they are intended as a universal system of "phonemic" representation, that is, a system for representing contrasts between the utterances of any language. Jakobson sought a feature system which was not only universal but also minimal; he thus presents many cross-linguistic complementary distribution arguments for unifying various different oppositions into a single feature, e.g., unifying the oppositions rounded/unrounded and pharyngealized/non-pharyngealized into the single feature /r/.

If a phonemic representation as envisioned by Jakobson were incorporated into a full grammar of a language, that is, a device which specifies how semantic representation is paired with phonetic representation in that language, that grammar would have to contain both a phonological component, which would assign a phonemic representation to the "surface syntactic" representation of each utterance, and a feature interpretation component, which would specify the relationship between phonemic representation and phonetic representation in that language. The feature interpretation component would presumably have to involve (1) rules predicting "non-distinctive" values of the universal features (e.g., rules which specify the non-distinctive voicelessness of final obstruents in Russian and the non-distinctive voicedness of obstruents which are followed by a voiced obstruent), (2) rules stating which of the possible realizations of each of the universal features is utilized in a given language (e.g., a rule of English that the feature of flatness is realized as lip-rounding rather than as pharyngealization), and (3) rules specifying "ideal values" for physical parameters involved in the realization of those features (e.g., a rule in Serbo-Croatian that a long vowel is about 1½ times as long as a short vowel, as opposed to Czech, where a long vowel is about 2½ times as long as a short vowel).

*This article appeared in French translation in *Langages* 8:112–123 (1967) under the title "Le rôle d’un système de traits phonologiques dans une théorie du langage." Reprinted by permission.
each morpheme is represented as a sequence of segments, each segment being a set of "underlying" specifications for phonological features. The "systematic phonetic" representation of an utterance is a representation in which each segment is assigned a value of + or − for each of the universal features, regardless of whether that value is "distinctive." The full grammar of a language would thus involve both a phonological component (which Halle describes in great detail) and a feature interpretation component (which he largely ignores), although the feature interpretation component would now only perform the second and third of the three functions listed above and would have for its input the "systematic phonetic" representation rather than a "phonemic" representation. The phonological rules operate in terms of the same system of features on which both the underlying representation of morphemes and the "systematic phonetic" representation of utterances are based. Each rule specifies the class of segments affected in terms of a formula involving those features and specifies its effect as certain changes in the feature composition of the affected segment, as in the Korean rule

\[ [+\text{obstruent}] \quad \text{[+closure]} \quad \text{[−tense]} \rightarrow \text{[+voiced]} \]

in env. \[ [+\text{voiced}] \quad \rightarrow \text{[+voiced]} \]

which changes to + the voicing specification of a non-tense stop or affricate which is preceded by a voiced segment and followed by a voiced segment, as in [pat] + [etta] → [padetta] ‘received’.2

Since Halle's purposes in Sound Pattern and subsequent works are so different from those of Jakobson in Preliminaries, the question obviously arises as to whether Jakobson's theories are adequate for Halle's purposes. It should first be noted that there are a couple of respects in which the feature system of Preliminaries is inadequate even for Jakobson's purposes, in that there are languages possessing contrasts which cannot be represented in that feature system. For example, the feature system of Preliminaries contains a feature of stridency, which is supposed to represent both the difference between affricates and ordinary stops and the difference between the "noisy" spirants [f, s, ŋ, x] and the "less noisy" spirants [φ, θ, s, ʃ]; however, the existence of languages such as Chipewyan, which has a three-way contrast between [t, tʰ, tʷ], shows that the difference between stops and affricates and the difference between "noisier" and "less noisy" places of articulation are two independent dimensions on which sounds may differ and thus may not be subsumed under a single feature, as in Preliminaries. Using the terms "abrupt release" and "proximal"3 to denote these two oppositions, the segments in question may be represented as

<table>
<thead>
<tr>
<th>Feature</th>
<th>Dental</th>
<th>Alveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closure</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Abrupt release</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Proximal</td>
<td>−</td>
<td>+</td>
</tr>
</tbody>
</table>

The above is an example of a case in which distinct articulatory oppositions which Jakobson subsumed under one feature may in fact not be so subsumed within an adequate universal theory of phonology. It is interesting to note that such may be the case even when the two articulatory oppositions never function as independent dimensions of contrast. Consider the feature of "flatness," under which Jakobson subsumes the articulatory oppositions of lip-rounding and pharyngealization. Arabic has an opposition between plain and pharyngealized consonants and has a three-vowel system [i, a, u], of which [u] is rounded; vowels are pharyngealized when adjacent to a pharyngealized consonant. Consider what phonological rules and "feature interpretation rules" would be needed to repre-

2In this paper I use square brackets to enclose all segmental transcriptions, phonetic or not, and use slashes // only to call special attention to the fact that a representation is the underlying representation. slashes here thus do not indicate "phonemic" representation, which never figures in my examples and which I indeed reject completely. Note that in a system of ordered rules the input to and output from a rule are generally neither "underlying" nor "systematic phonetic" representation but some intermediate stage of representation.

3These terms are introduced in Chomsky and Halle. The Sound Pattern of English (Harper & Row, to appear 1967). [See now Chomsky and Halle, 1968—VBM.]
sent these facts within a theory in which the phonological component of the grammar operates in terms of the flat/non-flat opposition and "rounded" and "pharyngealized" figures only in the feature interpretation component. The phonological component would have to contain the rule

\[
[+\text{syllabic}] \rightarrow [+\text{flat}] \text{ in env.}
\]

\[
[-\text{syllabic}]
\]

and the feature interpretation component would have to specify that "flat" is interpreted as "pharyngealized" when attached to a consonant or to a front or low vowel, as "rounded" when attached to a high back vowel, and as "rounded plus pharyngealized" when attached to a high back vowel which is adjacent to a [+flat] consonant. But note that this means that the generalization "vowel is pharyngealized when adjacent to pharyngealized consonant" must in effect be stated twice, once in the phonological component and once in the feature interpretation component. Thus, a theory of phonology in which pharyngealization and rounding are treated as the same feature throughout the phonological component suffers from exactly the same defect which Halle pointed out in theories which require underlying forms to be converted into phonetic representation through an intermediate stage of "taxonomic phonemic" representation: in each case the theory may force one to treat a single phonological process as if it were two rather unrelated processes in two separate components of the grammar. I accordingly conclude that even if there are no languages in which rounding and pharyngealization function as independent oppositions, a theory which treats them as separate must still be held superior to a theory which subsumes them under a single feature.

The above examples have related to the superiority of one feature system over another on the basis of the representations which it gives as output of the phonological component, in the one case because of the inability of one of the systems to distinguish things which contrast at that level, in the other case because the mode of representation forces the feature interpretation component to duplicate rules which are already part of the phonological component. However, a system of features may be superior to another on grounds relating exclusively to the phonological component, namely relating to the role which the features play in specifying the classes of segments to which the various rules apply and the effects which these segments are subjected to. An instructive example in this connection is the comparison of the features "compact" and "diffuse" with alternative features which could be employed to represent the same contrasts. The feature of "diffuseness" has been used to oppose high vowels and alveolar, dental, and labial consonants (which are all designated as [+diffuse]) to low and mid vowels and palatal and velar consonants (which are all [−diffuse]); the feature of compactness has been used to separate the low vowels ([−compact]) from the mid and high vowels ([+compact]); there has been no uniformity on the assignment of values for compactness to consonants, which in some works (The Sound Pattern of English) are all said to be [−compact], whereas in others (The Sound Pattern of Russian) the [+diffuse] consonants are said to be [−compact] and the [−diffuse] consonants [+compact]. Suppose that a theory of language involving the feature of diffuseness is compared with an otherwise identical theory in which diffuseness is replaced by a feature in which consonants are matched with vowels in precisely the opposite fashion, i.e., a theory which differs from the former by having instead of diffuseness a feature called "high" which opposes high vowels and velar and palatal consonants ([+high]) to mid and low vowels and alveolar, dental, and labial consonants ([−high]). Crucial for choosing between these two theories is a consideration of those phonological rules in the languages of the world in which either a class containing both consonants and vowels is involved in the rule or a vowel or consonant has an unambiguously "assimilative" effect on a consonant or vowel respectively: in the former case if the one feature system allows the class to be specified with a single feature value (say, [+high]), then the latter will require a disjunction involving another feature also (say, "diffuse consonantal or non-diffuse non-consonantal"), thus providing evidence that it is the features of the former system that are functional in the rule; in the latter case the effect of the rule will in the one system be to make

\[+-c\]

\[{}\]

\[\text{The Sound Pattern of Russian, pp. 22–23.}\]
the two segments agree in the value of the feature in question, in the other system to disagree, thus giving evidence that the former system more correctly characterizes the notion "assimilation."

In every case which I have been able to find which is of relevance to the choice between these two theories, evidence is provided for the superiority of "high" over "diffuse." (1) In Sanskrit [s] becomes retroflexed after [i, u, r, k]. Since Sanskrit [r] has palatal place of articulation and since palatal and velar consonants other than [k, r] become something else before [s] (i.e., all consonants become voiceless and unaspirated before [s] and palatal obstruents become velars), the rule can be expressed as

\[
\begin{array}{c}
\text{[+obstruent]} \\
\text{[grave]} \\
\text{[closure]} \\
\end{array} \quad \rightarrow \quad \text{[+high]}
\]

in env. [+high] ———

i.e., it is simply an assimilation of the feature of highness. If this rule were expressed in terms of diffuseness rather than highness, it would appear to do two quite distinct things: to assimilate diffuseness after a non-diffuse consonant and to dissipate diffuseness after a diffuse vowel. (2) In Maxakali, a language spoken in Brazil, there are phonological processes by which a vowel is added after a word-final stop and the stop is either weakened or deleted entirely. The vowels added are as follows: after [p], add [i]; after [t], add [a]; after [s], add [i]; after [k], add [i]. Note that the vowel added has the same "highness," i.e., the opposite diffuseness to the consonant to which it is added. Moreover, the choice of vowel is determined by an "assimilation" less movement is required of the organs of speech in passing from [p] to [i], etc., than would be required in passing to vowels of the opposite highness (i.e., the same diffuseness) as the consonant: [pi], [ti], [cai], [ko]. Thus the theory with highness instead of diffuseness correctly represents the assimilatory nature of this vowel insertion rule. (3) The extremely common phonological rule by which [iy] becomes [i] is universally classed as an assimilation and, in terms of the theory with highness, consists of the stop taking on the [+high] specification of the glide; however, in terms of the theory with diffuseness this change looks like a dissimilation: [t] and [y] are [+diffuse] but [s] is [−diffuse]. (4) In the Ripuarian dialect group of German, which includes the dialect of Köln, dentals have become velars after high vowels, as in [hunʃ] 'dog', [kitʃ] 'child', [luk] 'people', [tyk] 'time', corresponding to standard German Hund, Kind, Leute, Zeit. This change is also an assimilation of highness.

Not only are there cases such as the above which argue for classing high vowels together with velar and palatal rather than dental and labial consonants and to my knowledge no cases which argue for the opposite classification, but it is quite easy to give a uniform articulatory characterization of the [+high] segments: they are the segments whose primary constriction or closure is above a line drawn from the rear of the alveolar ridge to the uvula, whereas a uniform articulatory characterization of the "diffuse" segments may well be impossible; the articulatory definitions which have been proposed for them either do not define the class of segments enumerated above or do not provide a criterion which applies uniformly to consonants and vowels.

In this connection, it is worth bringing up the other feature in Preliminaries whose articulatory definition has never been satis

---

\textsuperscript{8}Data from V.M. Schlimmork, Deutsche Mundartkunde (Akademie Verlag, Berlin, 1962), p. 121.

\textsuperscript{9}I am grateful to Lester Rice for formulating this definition.

\textsuperscript{10}E.g., the definition in Halle's "In Defense of the Number Two" (in Studies Presented to Joshua Whitmore, Mouton, The Hague, 1957, pp. 63-72): "These two extremes of vocal tract shape, the horn and the Helmholtz resonator, are taken as the defining characteristics of the features compact-noncompact (horn shape or not), and diffuse-nondiffuse (Helmholtz resonator shape or not)." (Halle, p. 71.) This definition would make not only [p, t] but also [s] "diffuse."

\textsuperscript{11}E.g., the definition in Halle's "On the Bases of Phonology" (38) (in J. Katz and J. Fodor, eds., The Structure of Language, Prentice-Hall, Englewood Cliffs, 1964, pp. 324-333): "diffuse sounds are produced with a narrowing which in degree exceeds that of a constriction and is located in the front part of the vocal tract. . . . The dividing line between \textit{front} and \textit{back} is further retracted for vowels than for other sounds: for the vowels, \textit{front} includes almost the entire oral cavity, while for other sounds, the dividing line between \textit{front} and \textit{back} runs between the alveolar and palatal regions."
factory, namely "vocalic." The definitions given in "On the Bases of Phonology" ("vocalic" sounds are pronounced with a periodic excitation and with an open oral cavity, i.e., one in which the most extreme degree of narrowness is a constriction; "consonantal sounds are pronounced with occlusion or contact in the central path through the oral cavity") appear to preclude a segment from being both [+vocalic] and [+consonantal]; nevertheless, these two features are supposed to "produce a quadripartite division of the sounds of speech into (1) vowels, which are vocalic and non-consonantal; (2) liquids . . . , which are vocalic and consonantal; (3) consonants, which are non-vocalic and consonantal; and (4) glides . . . , which are non-vocalic and non-consonantal" (ibid., p. 327 [395]). One clear defect of the feature system of Preliminaries is that it provides no way of representing the difference between syllabic and non-syllabic liquid and nasals. Since there are languages in which utterances may differ solely by virtue of the syllabic or non-syllabic nature of a syllable (e.g., in many dialects of the verbal noun gambling has a non-syllabic [l] but the present participle gambling has a syllabic [l]), an adequate phonological theory must provide some feature to distinguish between these segments. Leaving aside the difficult question of giving an articulatory characterization of syllabicity, which I will denote this feature by, I note that it is the only systematic distinction between vowels and glides and moreover that there are many cases in which an alternation between vowel and glide is governed by exactly the same rule which governs an alternation between syllabic liquid or nasal and non-syllabic liquid or nasal (e.g., Sanskrit has a rule by which high vowels become glides and syllabic liquids become non-syllabic when a vowel follows). I accordingly propose to scrap the feature of "vocalic" and assert that the features "consonantal," "syllabic," and "obstruent" more adequately distinguish between the principal classes of segments:

<table>
<thead>
<tr>
<th>Vowels</th>
<th>Glides</th>
<th>Syllabic liquids</th>
<th>Nasals</th>
<th>Syllabic nasal liquids</th>
<th>Stops</th>
<th>Spirants</th>
<th>Affricates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocalic</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Consonantal</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Obstruent</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

This system of representation has several advantages over that in terms of "vocalic," among them (1) the fact that it makes non-syllabic liquids closer in feature composition to glides than to vowels (rather than vice versa, as in the system with "vocalic"), which fits well the fact that alternations between non-syllabic liquid and glide are extremely common but alternations between non-syllabic liquid and vowel are quite rare; and (2) in the extremely common case in which the underlying forms of the morphemes of a language are all of the form "CVCV . . . ," the system with "syllabic" allows that generalization to be reflected in a redundancy rule which specifies every odd-numbered segment as [-syllabic] and every even-numbered segment as [+syllabic], whereas in the system with "vocalic," all that could be predicted about vocalicess is that even numbered segments are [+vocalic] and odd-numbered [-consonantal] segments are [-vocalic], i.e., the generalization about morpheme shape would make the vocalic specification of glides redundant but would not make redundant any feature specification in a liquid, nasal, or obstruent.

It should be clear from the preceding portions of this paper that the optimum feature system for a theory of language such as that of The Sound Pattern of Russian and subsequent works will involve a considerably greater number of features than appear in Preliminaries and Sound Pattern of Russian. Indeed, on the basis of arguments such as those which I have presented above,11 Chomsky and Halle conclude in Sound Pattern of English that a system containing over twice as many features as that of Preliminaries is needed to provide for adequate phonological description. Since the values of n binary features may be combined in 2^n different ways, the question immediately arises of whether there are as many segment types as can in principle be represented by a system of features such as that of Sound Pattern of English, i.e., 2^m = 29,434,432. Clearly there are not: there are very sharp constraints on the way that feature specifications may combine, and the number of combinatorial possibilities is of the thousand, even that of the phonological feature combination larger than the phonetic representation as a fact of the fact that phonologists are met with representations that play a role in normal coding; we expect it to be.

In actual fact, the feature combinations are not an important fact of phonology but that the universal combinations provide a phonological rule mentioned in in a feature specification that a feature specification is adjusted for feature specification feature specifications. For example, [released] [-aspirated] [-released], the rule also makes [-aspirated] make it [+released] the important role in phonological facts from Korea.

In syllable only segments [l], nasals, and ing to this fact, neutralized in a syllable-final position /nač/ 'day', which are distinguished [nače], [naće], nominative: [nač].

11For example Some Recent Ch. [41].
of combinations which will appear in the "systematic phonetic representations" of natural languages is probably no more than a couple of thousand. The fact that the feature systems (even that of Preliminaries) of generative phonological works allow in principle a class of feature combinations which is astronomically larger than the class which will appear in phonetic representations has been taken by some as a defect of generative phonology. However, this fact merely shows that the goals of the phonologist and the electrical engineer cannot be met within the same representation: a representation in terms of the categories which play a role in phonology will not be an "optimal coding"; of course, there is no reason to expect it to be.

In actual fact, the restrictions of feature combinations which I mentioned above play an important role in the functioning of the phonological component of a language. Briefly, the universal constraints on feature combinations provide a mechanism whereby a phonological rule may have effects other than those mentioned in the rule: when a rule introduces a feature specification into a segment, the segment is adjusted so as to acquire all other feature specifications which are implied by that feature specification and the universal constraints. For example, there is a universal constraint excluding the combination of features [+released] and [+aspirated]. When a rule makes a segment [+aspirated], that segment is automatically also made [+aspirated]; likewise, a rule which makes a segment [+aspirated] would also make it [+released]. An excellent example of the important role which this principle may play in phonology is given by the following data from Korean.

In syllable-final position in Korean the only segments which occur phonetically are [l], nasals, and unreleased stops. Corresponding to this fact, many underlying contrasts are neutralized when a consonant comes to be in syllable-final position; for example, underlying /naː/ ‘day’, /naːβ/ ‘face’, and /naː/ ‘sickle’, which are distinct in the locative: [naje], [naβe], [nase], become homophonous in the nominative: [nɑː] (where - means "unreleased"). In this alternation, underlying obstruents become unaspirated and lax, and in addition, underlying /s, ç, h/ all become [l]. There are universal constraints excluding the feature combinations [+aspirated] and [+released]. Consequently, a rule which made consonants in syllable-final position [+released] would automatically make them also [+aspirated] and [+closure]. Thus the effect of the rule

[±syllabic] → [±released]

in env.

{[±syllabic]}

(# means word-boundary) would be to turn [s, ç, h] into [l-, l-, ?] respectively, where [l-] denotes an unreleased apico-palatal stop. The eventual effect is obtained as the result of subsequent rules which turn glottal stop into an apical stop and which make unreleased apical stops dental:

[±consonantal] → [±grave]

[±released] → [±high]

Korean is traditionally regarded as having a single underlying liquid, which is pronounced as flapped [r] intervocally and as [l] in syllable-final position or in a gerund; an underlying liquid at the beginning of a word is either deleted or turned into [n], depending on whether or not it is palatalized. Examples: [kʰal] ‘knife (nom.)’, [kʰare] ‘knife (loc.)’; [təro] ‘highway’, [nobyen] ‘roadside’; [itsu] ‘mileage’, [oryi] ‘five miles’. There are not only verbs in which morpheme-final [r] alternates with [l]: [alg] ‘to know and’, [aretta] ‘knew’, but also a large class of verbs in which morpheme-final [r] alternates with [l]: [mutkʰa] ‘to inquire and’, [muretta] ‘inquired’. The latter verbs are traditionally regarded as irregular; however, the rules given already plus the universal constraints make it possible to assign to these verbs an underlying representation such that these alternations automatically arise. Specifically, note that [l-] is what would arise by the universal constraints from an underlying flapped [r] which was made unreleased. Thus, if Korean is analyzed as having two distinct underlying liquids, [l] and [r], and the two verbs are represented as underlying [al] and
\(/mur\), the rules given already would yield all the desired forms. Note further that because of the rules alluded to whereby an intervocalic liquid becomes \(/r\) and a word-initial liquid either is deleted or turns into \(/n\), morpheme-final position is the only underlying segment position in which an underlying distinction between \(/l/\) and \(/r/\) could ever have any phonological effect. Since an underlying feature specification for laterality,\(^{13}\) which distinguishes \(/l/\) from \(/r/\), plays no role in the phonological rules of Korean unless it is in morpheme-final position, all occurrences of liquids in other positions may be left unspecified for laterality in their underlying forms, and the value realized phonetically would be inserted by one of the phonological rules just mentioned. Consider now the fact that not only are there verbs which display the normal alternation between syllable-final \(/p\) and intervocalic \(/b\): \(\text{[cep]}'\) 'to fold', \(\text{[cebe]}'\) 'to fold and', but there are also verbs which display an alternation between \(/p\) and \(/w\): \(\text{[kip]}'\) 'to mend', \(\text{[kiwe]}'\) 'to mend and'. Here again underlying forms can be set up which will automatically yield the correct results thanks to the universal constraints: \(\text{[cep]}/\text{[kiwe]}\). When it is made unreleased, underlying \(/w/\) becomes a stop as a result of the universal constraints.

An interesting point to note concerning the above discussion is that the realization of Korean which I propose requires that "released" play a role within the phonological component. This fact is noteworthy since no cases have been reported of a language in which the distinction between released and unreleased stops is distinctive. I point this out in order to emphasize that the contrast of features which play a role within the phonological component is anything but the extremely limited class of largely "distinctive" features which it until recently was generally assumed to be, and that the phonological component, rather than affecting the "more distinctive" features in the "earlier" rules and the "less distinctive" features in the "later" rules, as is sometimes supposed, must operate in terms of highly "non-distinctive" features even in very early rules of the grammar, such as the rule making syllable-final consonants unreleased indeed is.

In conclusion, it would be worthwhile to reconsider for a moment the role which the "feature interpretation component" plays in a grammar. Of the three functions which it would play in a grammar containing a "taxonomic phonemic" level of representation, one becomes unnecessary if that level is rejected. Of the two remaining functions, the domain of one, namely that of choosing between alternative modes in which a feature may be realized, has been reduced considerably in the process of the above arguments. Indeed, there remain no clear cases in which anything is gained by considering two distinct articulatory features to be identified with the same feature of universal phonology. Accordingly, then, the task of the feature interpretation component of a grammar can probably be restricted to that of supplying ideal values to the physical parameters in which the features which function in the phonological component manifest themselves.