SCAMP Working Paper No. 7/81

ON THE RELATIONSHIP OF PHONOLOGIC FEATURES TO PHONETIC PARAMETERS

Morris Halle

ABSTRACT

It is argued that the point of view that a close relationship obtains between phonologic and phonetic features is correct. A statement of an opposing point of view made by Peter Ladefoged in his presidential address to the Linguistic Society of America is subjected to close scrutiny and arguments are advanced to demonstrate its weaknesses. An attempt is made to identify the essential differences between the two points of view: they appear to agree on the question of requiring some internal representation of words, but disagree on the question of the relationship of this representation to the output stage, Halle arguing that the output representation is formally close to the form in which the words are represented in memory and asserting that Ladefoged would deny that there is any principled basis for such an assertion. The point of view defended by Halle is that, in principle, there cannot be a natural language for which the internal representation of words is as unrelated to words as the productions of a trained touch typist appear to be.

F7

Phonetics is the study of speech, i.e., of the physical manifestation of natural language. It is, therefore, concerned with the articulatory gymnastics that goes on in our vocal tract when we speak, and with the acoustical signal that is the end product of this gymnastics. Much that we know about these matters has been learned from direct observation of phenomena. Extensive studies of the articulatory process have been conducted with the help of x-rays, fiber optics, and other devices that allow close observation of the activities in the vocal tract during speech. Similarly there have been numerous investigations of speech as an acoustical signal: all of us here have had occasion to study speech with the sound spectrograph and other kinds of electro-acoustical analyses. While such studies have contributed much of great value to our discipline, they have been unable to answer basic questions about how speech is actually produced and processed by human subjects. The reason for this, it seems to me, is

not at all obscure: we have not learned all we might wish to learn from such studies because we have arbitrarily restricted our concern to factors that are accessible to direct observation and have excluded from discussion equally pertinent factors which differ from the former only in that their effects are not directly observable in the stimulus but must be inferred by more-or-less complex chains of reasoning.

Perhaps the single most important factor that this self-denying ordinance against abstractness would eliminate from the purview of our science is the subjects' knowledge of the language in which the stimuli under study are framed. Under normal circumstances, this knowledge plays a fundamental role both in the production of an utterance as well as in its perception. With regard to speech production this fact was remarked upon long ago in a classical paper by Sapir who pointed out that the articulatory gesture involved in blowing out a candle is fundamentally different from the superficially identical gesture involved in the production of the voiceless glide in _when_, _whisky_, _wheel_. With respect to speech perception we can cite the recent paper by Remez _et al_. (1981) where it was shown that an acoustical signal, which under one set of instructional

conditions is perceived as computer bleeps or science fiction sounds, is perceived as a specific English utterance when the instructional conditions are different. It hardly needs stressing that this result crucially depends on the fact that the subjects in this experiment knew English. Compared to this factor, the precise characteristics of the physical stimulus were of considerably lesser importance.

If knowledge of language is to be included in the study of speech it is appropriate to inquire how one might go about obtaining information about the character of this knowledge. Given the present status of science it is fair to assume that for a considerable time into the future it will be impossible to establish what speakers know about a language with the help of electrophysiological recordings or other means of direct observation of their central nervous system! This, however, does not mean that we are, therefore, condemned to ignorance in this domain. Physicists have for generations learned much about the nature of forces and particles which they cannot ever hope to observe directly. They have done this with the help of the hypothetico-deductive method which is at the heart of all science. Reduced to its essentials the method consists

of the formulation of a highly articulated theory which makes detailed claims about observable properties of the phenomena under study. The complexity of the chain of reasoning connecting the theoretical entities of a science to the directly observable phenomena has never been shown to affect the validity of the theoretical constructs, provided, of course, that the chain of reasoning was sound, and I know of no arguments that special conditions hold in the case of speech that would justify restricting theory construction in this domain in special ways. I conclude, therefore, that in the study of speech the hypothetico-deductive method, standard in other branches of science, can and must be used.

This conclusion is, of course, far from revolutionary; in their studies of speech, many linguists and phoneticians have taken full account of the role that knowledge of language plays in this domain. As an example, consider the contributions to our understanding of speech that are due to the leaders of the Prague school, Roman Jakobson and Prince N. S. Trubetzkoy. It was they who showed that, in many cases, the classes of speech sounds that arise when sounds are grouped together by virtue of being treated identically in one or more rules of a language are coextensive

with the classes that arise when speech sounds are grouped together on the basis of shared phonetic properties, such as lowering of the velum, vocal vibration, lip rounding, etc. For example, it is well known that the basic rule of plural formation of English nouns groups the words into three classes depending on the word-final sound. As illustrated below, when the word ends with the sounds [s z č ǰ š ž], the plural suffix is [əz]; when it ends with [p t k f θ], the suffix is [s]; otherwise, the suffix is [z].

(a) taxes roses leeches bridges
sashes garages

(b) troops trouts tacks coughs
sixths

(c) cobs cords cogs coves lathes
rims tins songs bars balls
spas spies pillows cues cows.

In the first group are words that end with what phoneticians would describe as coronal affricates and fricatives; in the second group, the nouns end with voiceless ~~sounds~~ consonants; and in the third group, the nouns end with any sound but those just mentioned. The plural rule of English might, therefore, be stated as in the following:

Plural suffix is

(a) [əz], if noun ends with coronal strident consonants; otherwise,

(b) [s], if noun ends with voiceless consonants; otherwise

(c) [z].

The observation that the classes of speech sounds, which underlie the different rules and regularities in the language, are phonetically natural in the sense just illustrated led to the further hypothesis that a single system of parameters underlies both the overt manifestation of language in speech and the more covert yet equally real rules and regularities of language, and that one could and should make use of information about language--that is, of information from rules and other regularities--in searching for the parameters that govern speech.

This view of the close relationship between phonologic and phonetic features has not enjoyed universal approbation. Unfortunately, until relatively recently there was in the literature no detailed statement of the contrary position. This made discussion difficult since defenders of the close relationship between phonology and phonetics were forced to argue against views that had not been stated

explicitly by anyone, and therefore, were often in danger of arguing against straw men of their own creation. This obstacle to progress has now been removed, thanks to Peter Ladefoged, who, in his presidential address to the Linguistic Society of America in 1978 and in a number of other papers, has articulated the opposing position. Since the only one of these papers that was accessible to me last week was Ladefoged's presidential address, I shall concentrate on that paper here. Ladefoged summarizes his views in the concluding paragraph of his address as follows:

> In summary, I have tried to show that the fundamental linguistic phonetic constraints are sets of articulatory and acoustic parameters. Each set is a necessary and sufficient set of parameters that will account for all possible linguistic phonetic properties. Descriptions in terms of one set can be converted into descriptions in terms of the other. Descriptions of phonologic patterns in language involve features which are quite distinct from the phonetic parameters; moreover they cannot account for many of the phonetic differences between languages. At some abstract level languages may be organized partly in terms of phonological features. But we must always remember that languages are complex properties of human societies, not of individual brains. Individuals producing and interpreting linguistic events probably use something like the [articulatory and acoustic - MH] parameters in Lists 1 - 2 [of the paper] (pp. 501-2).

There are thus, according to Ladefoged, two kinds of entities. On the one hand, there are articulatory and acoustic parameters which are used by individuals in producing and interpreting utterances. These parameters have, therefore, psychological reality. The phonological features, on the other hand, are entities distinct and separate from the phonetic parameters. Since phonological features are properties of language and since "languages are complex properties of human societies, not of individual brains" it would appear that, in Ladefoged's view, phonological features differ from phonetic parameters in that they lack psychological reality. This impression is further strengthened by Ladefoged's comment elsewhere in the paper that phonological features "have in no way proved to be the mental representations used by people when speaking and listening to language. Most of them are completely unnecessary for the adequate description of the behavior of speakers and listeners" (p. 496). Having thus denied that features are a part of the mental representations of speakers and listeners, Ladefoged continues somewhat unexpectedly: "But if they are mental representations, then I would like to know what they are mental representations of. The best answer that I can

come up with is that they are part of the mental representation of what a speaker knows about the social institution called language" (p. 496). If the phonological features are among our mental representations, then regardless of what they represent, they must be in the human brain, for the brain is the repository of all mental representations in our species. But if this is the case, then we are confronted with the need to reconcile this fact with Ladefoged's remark, quoted earlier, that features are not properties of the human brain. This can be done only by disregarding one or the other of the two statements. We can either assume that Ladefoged does not really wish to deny mental representation and psychological reality to phonological features. Alternatively, we can disregard the statement about the features being "part of the mental representations of what a speaker knows about the social institution called language." Since it is rather implausible that anyone would deny that knowledge of language is part of the mental equipment of normal speakers, I shall choose the former alternative and disregard the remark about features not being properties of speakers' brains.

This assumption has also the virtue of narrowing the issues in dispute. Since the psychological reality of features has been eliminated as an issue, the disagreement now concerns the remoteness of the relationship that holds between phonological features and phonetic parameters. In Ladefoged's view the two sets are remotely related; in the more traditional view that I am defending here, the relationship is quite close. Ladefoged does not characterize the relationship in great detail. He observes that "the properties of the abstract thing we call language are different from those required in a set of rules required for sentence perception" (p. 496). He provides several examples in the paper where the relationship is quite indirect. For instance, he points out that his articulatory parameter of back-raising "is not very useful in explaining observed vowel patterns, or in writing phonological rules for alternations of vowels" (p. 490). By contrast, in the case of the articulatory parameter of front-raising, the relationship to the phonological feature [± back] is quite direct, for as Ladefoged notes, this articulatory parameter "clearly separates front vowels from back vowels" (p. 489).

To get an idea of what a remote relationship might be like, it seems worthwhile to digress momentarily in order to consider the relationship that obtains between the phonologic features of English and the parameters that enter into the generation of the letters of our cursive script. The letters of our cursive script, as those of us who were taught the Palmer method of penmanship in grade school will no doubt remember, are composed of elementary strokes as illustrated in Figure 1.

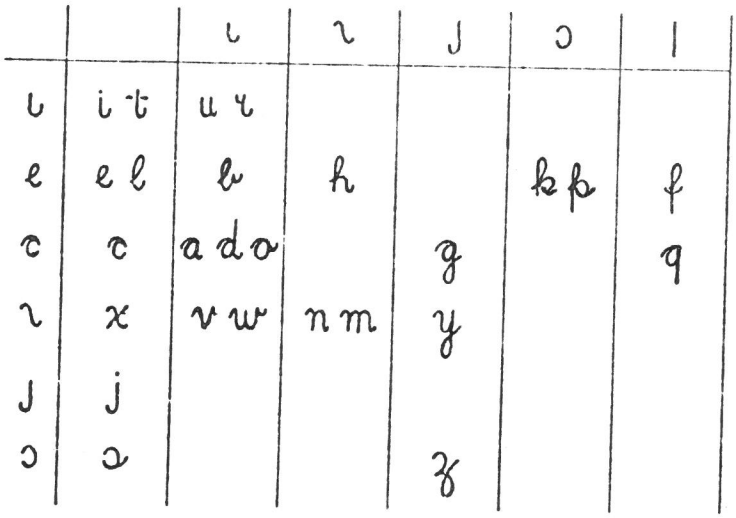

Figure 1. Elementary strokes used in writing lower-case alphabet.

It is obvious that letters sharing a particular graphic property--e.g., the e-loop or the c-hook--do not represent sounds that share either articulatory or phonologic properties. Although we do not have generally recognized criteria for measuring closeness or remoteness between different sets of parameters, it is reasonably obvious that the relationship between the graphic parameters of the letters in our cursive script and the phonological features of the sounds represented by these letters is considerably more remote than that between the phonological features and Ladefoged's articulatory parameters. In the latter case, there are a number of instances where features and parameters stand in a direct one-to-one relationship. We have noted this already with regard to the parameter front-raising and the feature [± back]. It appears also to be true with respect to the parameter velic opening and the feature [± nasal], and, as Ladefoged notes (p. 492), the parameters of pharynx width and tongue bunching "correlate in a fairly simple way with phonological features." Since there are no examples of such direct relationships between graphic parameters and the phonological features, we conclude that the graphic parameters and the phonological features are unrelated.

The question that arises at this point is whether the fact that the classes formed by certain of the articulatory parameters coincide precisely with those formed by certain phonological features is a mere accident, a fortuitous parallelism with no further significance, or whether this close parallelism reflects systematic regularities inherent in the structure of language. If the latter is the case--as it has been assumed by many scholars--then what has to be explained is not the instances where there is a correlation between the two sets of properties, but rather parameters such as Ladefoged's back-raising that are "not very useful in explaining observed vowel patterns, or in writing phonological rules for alternations of vowels" (p. 490). Since articulatory parameters are not God-given, but are rather suggestions made by error-prone humans, the hypothesis must be seriously entertained that Ladefoged's parameters which lack a direct relationship to phonological features are, like phlogiston in chemistry, the inheritance of acquired characteristics in biology, and the bi-unique phoneme in linguistics, theoretical postulates that have failed scientific scrutiny and are, therefore, destined to disappear from scientific discourse sooner or later, to be replaced by parameters that correspond more closely to phonological features.

In order to estimate the likelihood of such changes in the list of parameters it is necessary to examine the basis on which the parameters were included in Ladefoged's list. Ladefoged's articulatory parameters derive from those of Coker, Umeda and Browman (1973) who "showed that it is possible to use articulatory specifications to produce intelligible English. The input to their computer program was a string of phonetic segments that were changed by the program into ten articulatory parameters. Insofar as the sounds produced were like English, these parameters were sufficient to specify the sounds involved." Ladefoged found that when a wider range of languages is considered the number of parameters must be increased; however, these additional parameters are selected on the same basis as the original ten; i.e., they are sufficient to specify the articulatory configurations that produce the appropriate acoustic output.

The fact that the parameters are sufficient to specify articulatory configurations that result in the appropriate acoustic output is not of itself sufficient to guarantee that these parameters are the ones used by normal speakers. Because there are many alternative ways of achieving a particular acoustic effect it is, in principle, conceivable

that the required acoustic effects in a particular case are attributed to changes in the vocal-tract geometry that are implausible or even downright impossible from an anatomical point of view. Ladefoged is aware of this danger. At the end of the detailed discussion of the nature and function of front-raising and back-raising, the two major articulatory parameters controlling the shape of the tongue, Ladefoged writes: "We must now consider whether descriptions of the body of the tongue in terms of front-raising and back-raising parameters are simply mathematical abstractions, or whether they can really help us explain why vowels are as they are." Ladefoged clearly believes that his articulatory parameters are more than mathematical abstractions, and he states his reasons in the following passage, which because of its importance to this discussion is quoted verbatim:

> It seems, in fact, as if they [front-raising and back-raising - MH] might well summarize some of the principal muscular forces involved. The tongue and mandible form a very complex system, with a wide variety of potential actions (Hardcastle, 1976; Lieberman, 1977). As may be seen from Figure 4, the front-raising parameter corresponds in great part to the actions of the genioglossus, and of opposing muscles such as the glossopharyngeus and other pharyngeal constrictors. The back-raising lowering parameter effectively summarizes the opposing actions of the styloglossus and the hyoglossus. However, there

> are many possible compensatory actions of the jaw and the tongue muscles, and it is probably not too profitable to consider either parameter as simply specifying the action of a group of muscles. It seems more likely that the parameters (and perhaps others that I will be discussing) describe higher-level cortical control functions. That is, we may think of them as underlying parameters that determine the synergistic actions which are required for the skilled motor movements that occur in speech (pp. 488-9).

This passage does not constitute an argument. At best, it is a not-implausible speculation. Ladefoged, no doubt, is correct in attempting to motivate his articulatory parameters by appealing to anatomical and neurological evidence. Unfortunately, at present we know so little about these matters that no useful facts in its support could be cited (by Ladefoged or anybody else). In the absence of such evidence, Ladefoged's articulatory parameters must remain convenient mathematical abstractions with no claims to preference over many other equally convenient mathematical constructs. The search for relevant anatomical and neurological evidence must, and obviously will, continue. The review of the state of research on movement and other actions of organisms contained in Gallistel's (1980) recent book suggests that we may be on the threshold of genuinely exciting advances in this area. When these advances come

they are likely to revolutionize our conception of the speaking process and with it also the list of articulatory parameters. This constitutes yet another reason why it would seem advisable at this juncture to reserve judgment on how remote the relationship is that obtains between articulatory parameters and phonological features. One final reason for caution here is the fact that the phonological features are not cast in concrete either and may be revised in the next few years.

To summarize, no conclusive case was made, in the article under discusion, for the claim that the articulatory parameters are only remotely related to the phonological features. We found that among the articulatory parameters in Ladefoged's list there were several that were most directly related to phonological features. We also saw that the articulatory parameters in the list were motivated only insofar as that they permitted specification of articulatory configurations having appropriate acoustical outputs; i.e., they were the best outcome that could be obtained by curve fitting. There is reason, therefore, to suppose that, as the scope and variety of data under study increases, the list of parameters will undergo not inconsiderable changes in the future. It is not implausible to expect that, as

the parameters evolve, all of them will come to resemble such parameters as velic opening or front-raising, which, as we have seen, correspond directly to phonological features. Surely there is nothing in the paper or elsewhere in the literature suggesting that this is an unlikely direction for phonetic science to evolve in. Finally, the paper did not show that there are fundamental differences between articulatory parameters and phonological features with respect to their mental representations or psychological reality.

It seems to me that, at the heart of the controversy over the directness of the relationship between phonetic parameters and phonological features, there lie differences in the conception of the speaking process. Such differences are rarely discussed because they are themselves neither directly related to objective observations, nor are they readily resolvable into logical or mathematical computations. They are, if you will, the preconceptions and prejudices that each of us brings to our work and that account for the fact that different scientists approach a given problem in ways that are often fundamentally different. I believe that such a conceptual difference lies at the heart of the present controversy and I shall try in

this concluding section of the paper to make this difference explicit.

Consider an English-speaking phonetician making a phonetic transcription of an English utterance. It is immaterial to our discussion here whether the utterance is being dictated to the phonetician or whether the phonetician is writing out an original statement in IPA notation. To make matters particularly simple, assume that the phonetician uses an ordinary typewriter to make the transcription. Given these conditions we know that, at the output stage, the information that the phonetician processes must be in the form of signals to the muscles controlling the positions of the phonetician's ten fingers. We know also that this last stage is very remotely related to any other representation of the utterance being transcribed, since the output stage here is determined basically by the arrangement of letters on the typewriter keyboard which is the result of a historic accident of little linguistic or phonetic significance.

Consider now the situation that obtains when the same phonetician produces the utterance phonetically. The output stage here consists of signals to the anatomical structures of the phonetician's vocal tract. We also know

that the words that make up the utterance the phonetician is producing are part of the phonetician's vocabulary, for the phonetician can answer all sorts of questions and perform various sorts of transformations on the utterance, and this can be done only if the words are in the phonetician's vocabulary; i.e., if they are stored in some form in the phonetician's memory. The question that Ladefoged and I disagree about is the relationship of this representation to that of the output stage in the two cases. Whereas we agree that the internal representation of words, whatever its precise character, is quite unrelated to the output stage in the case of the typewriting example, we disagree about the nature of the relationship in the case where the output is speech. I would guess that, in this case, the output representation is formally close to the form in which the words are represented in memory. Ladefoged would deny that there is any principled basis for such a guess. Notice that I am not attributing to Ladefoged the absurd view that, in the two cases, there is no difference in the relationship between the internal representation and the output stages. We agree that there is a difference. What we disagree on is whether the difference is one of principle. In my conception there could not be a natural

language where the internal representation of words was as unrelated to the articulatory output as it is in the typewriting example; in Ladefoged's conception of language this is a possible--though perhaps not very likely--state of affairs.

The results of the research of the next few years are likely to decide between the two differing conceptions of language.

REFERENCES

Coker, C.H., Umeda, N. and Browman, C.P. Automatic synthesis from ordinary English text. IEEE Trans. Audio Electroacoust. AU-21, 1973, 293-298.

Gallistel, C.R. The organization of action (Hillsdale, NJ: Erlbaum, 1980).

Ladefoged, P. What are linguistic sounds made of? Language, 56, 1980, 485-502.

Remez, R.E., Rubin, P.E., Pisoni, D.B. and Carrell, T.D. Speech perception without traditional speech cues. Science, 212, 22 May 1981, 947-950.