Preliminaries to linguistic phonetics. By PETER LADEFOGED. Chicago: University of Chicago Press, 1971. Pp. xii, 122.

Reviewed by MORRIS HALLE, *MIT**

Perhaps the most welcome change to be observed in the field of phonetics as it has evolved during the last decade is the increasing attention being devoted on all sides to the elaboration and justification of its theoretical framework. The time when phoneticians were mainly interested in gathering curious facts without much thought about the significance of their findings, and when serious linguists could justifiably liken phonetics in its scientific status to numismatics and philately, has now passed, and few mourn its passing. The book under review is a good example of this new trend. Its aim is to establish a universal-feature framework for the characterization of speech sounds; this theoretical aim governs everything that appears in the book. As a result, the book concentrates not on piling up vast bodies of data but rather on constructing arguments, and on using the data in support of various theoretical claims. This fact naturally determines the major outlines of my review: it focuses of necessity on the claims made by Ladefoged and on the arguments marshaled in their support. Since it is clearly impossible to deal in a review with every issue of interest raised even in such a relatively short book as this one, I have limited the discussion below to four topics which seem to me to need further discussion. These four topics have been chosen both because of their intrinsic importance and because of doubts which I have with regard to their treatment by L. In the case of the topics discussed in §§ 1 and 2 below, L's arguments are clear and to the point, but fail to convince me because important evidence bearing on the issues has been overlooked. In the case of the topics discussed in §§ 3 and 4, the difficulty appears to me to be due rather to the arguments themselves: they do not seem to me to support L's claims.

1. For many years, voicing and aspiration appeared to be among the most solidly supported features of the phonetic framework. This situation has changed rather noticeably in the last few years, especially as a result of a series of studies by Lisker & Abramson (e.g. 1964), where questions were raised concerning these features. As expected, the current debate is directly reflected in L's book. He does not simply repeat the timeworn formulas about the nature of these features, but strikes out in a new direction.

According to L, 'in the formation of voiced sounds the vocal cords are adjusted so that they are almost touching along their entire length' (7), whereas in voiceless sounds 'even the anterior part of the glottis is so far apart that it cannot be set in vibration' (18).[1] But

* This work was supported in part by grants from the National Institutes of Health (5 TO1 HD00111) and the National Institute of Mental Health (5 PO1 MH13390).

[1] L also states (9) that 'during voiceless sounds the vocal cords are apart at the posterior end between the arytenoid cartilages (see Fig. 2b).' This statement is obviously incompatible with the description of voicelessness quoted above. It would seem that the description on p. 9 is in error: on the one hand, it does not describe the vocal-cord configuration shown

these two configurations are not the only ones that play a role in language. L notes that Gujarati and some other languages possess a set of vowels 'distinguished by a different adjustment of the vocal cords in which the posterior portions (between the arytenoid cartilages) are held apart, while the ligamental parts are allowed to vibrate' (12). L designates this type of phonation by the term 'murmur' ('breathy voice' is another common term), and states that it is also present in stop sounds such as the Sanskrit or Hindi [b^h d^h g^h]. To deal with these facts, L postulates a feature of GLOTTAL STRICTURE, which can assume at least three distinct values: voiced (vocal cords almost touching along their entire length), murmur (ligamental portion of vocal cords touching, arytenoid portion held apart), voiceless (vocal cords held apart along their entire length). In addition, L recognizes (following Lisker & Abramson) a feature of VOICE ONSET, which also can assume at least three distinct values, according to whether the onset of voicing precedes, approximately coincides with, or follows with a substantial lag the release of the consonantal stricture. A stop system like that of Hindi is, therefore, characterized in terms of the proposed features as follows (cf. p. 97):

	*p*ʰ	*p*	*b*ʰ	*b*
Glottal stricture	0	0	1	2
Voice onset	2	1	2	0

It is clear that the gross phonetic facts can be described in such a manner. However, as L himself notes, this is far from sufficient: 'Accounting for systematic phonetic contrasts is in itself a trivial and uninteresting task that can be done in many ways. It becomes interesting only when we try to constrain our account so that it fits in with the division of sounds into the natural classes required in phonological rules' (4). L therefore cites a number of phonological facts that he believes support his proposed features: 'Murmured or breathy voiced sounds are between voiced and voiceless sounds, and hence can be grouped with either of them; this is as it should be for appropriate descriptions of languages such as Shona and Punjabi. Similarly, voiced sounds and different forms of laryngealized sounds are a more closely related natural class than laryngealized and voiceless sounds, which is what is required in descriptions of Kumam. Furthermore, this formulation assists us in making statements about co-articulated allophones' (19).[2]

In view of the importance which L rightly attaches to linguistic data as corroborative evidence for his proposal, it is surprising that he does not bring up here what is no doubt the best known phonological rule involving the features under discussion, i.e. Grassmann's Law for Greek and Sanskrit, which in fact casts doubts on the adequacy of the proposed feature framework. Given the traditional features, the rule is stated as follows:

(1) [−continuant] → [−aspirated] / ___V [+aspirated]

That is, stops are unaspirated if followed in the same stem by an aspirated stop. In the

---

in Fig. 2b, and, on the other hand, it fails to distinguish the configuration for 'voicelessness' from 'murmur', described as being produced with 'the posterior portions (between the arytenoid cartilages) ... held apart, while the ligamental parts [of the glottis] are allowed to vibrate' (2).

J. T. Ritter has drawn my attention to certain facts which cast doubt on L's assertion that 'murmured or breathy voiced sounds are between voiced and voiceless sounds'. It is a well-known fact that voiceless sounds tend to cause tone raising in adjacent vowels, whereas voiced sounds tend to cause tone lowering. (Cf. Haudricourt 1961, where this is extensively documented for Southeast Asian languages.) Thus, if murmured sounds were indeed intermediate between voiced and voiceless, one would expect that the tone in an adjacent vowel would be raised more by a murmured sound than by a voiced sound, and one would not expect that the tone would be lowered more by a murmured sound than by a voiced sound. As a matter of fact, according to Cope 1970, it is the latter ('unexpected') case that obtains in Zulu; i.e., it is the murmured sonorants and obstruents rather than their voiced cognates which act as 'depressors' on the adjacent vowel tones. This is yet another fact that makes me doubt the correctness of L's treatment of the two features under discussion.

framework of L's book, the rule must be stated thus:

$$(2)\ [1\ \text{stop}] \rightarrow \left\{ \begin{array}{l} [1\ \text{Voice onset}] \,/\, [\ \underline{\quad\quad\quad}\ \ 0\ \text{Glottal stricture}] \\ \left[\begin{array}{l} 2\ \text{Glottal stricture} \\ 0\ \text{Voice onset} \end{array}\right] / [\ \underline{\quad\quad\quad}\ \ 1\ \text{Glottal stricture}] \end{array} \right\} \text{V}\ [2\ \text{Voice onset}]$$

That is, voiceless stops are replaced by voiceless stops having a voice onset which coincides with the stop release, while murmured stops are replaced by voiced stops with voice onset preceding the release, if followed in the same stem by a sound with voice onset that follows the release.

It is obvious, from a comparison of 1 and 2, that the traditional features are well suited to express Grassmann's Law, whereas L's proposed features are not. Since natural phonological processes, and only such processes, should be elegantly expressed in a good feature framework, the issue turns on whether Grassmann's Law does or does not represent a natural phonological process—which, according to most scholars, it does. It is to be regretted that L did not see fit to discuss Grassmann's Law here: until we know his reasons for disregarding this evidence, we are not in a position to accept his proposed features as an improvement over the traditional set.[3]

2. In the feature frameworks that have been most widely accepted, e.g. that of the IPA, consonants are characterized by a multi-valued feature of ARTICULATORY PLACE, which is not utilized in the characterization of the vowels. Vowels, on the other hand, are characterized by the features of BACKNESS and HEIGHT, which are not utilized in the characterization of the consonants. L adopts these three features (in somewhat modified form), and notes that the traditional restriction of the features to mutually exclusive classes of speech sounds makes it difficult to state certain natural phonological processes. He proposes to overcome this difficulty by supplying characterizations for both types of sound in both sets of features: 'In the feature system being proposed here, vowels will be assigned a value both for the backness and height features and also for the articulatory place feature; and consonants will have values for backness [and height] as well as place' (80).

One may view this dual specification of speech sounds in one of two ways. On the one hand, one may regard it simply as a programming trick, where a name is assigned to an arbitrary class of entities in order to facilitate some computation. Such a trick, however, involves no empirical claim about phonetics, and hence has no place in a book such as this. The dual specification of the vowels and the consonants must, therefore, be understood as implying a claim about the character of these sounds, specifically that vowels are to be characterized not only in terms of backness and height, but also in terms of articulatory place. Since no limitation on the co-occurrence of the features under discussion is stated, the reader can only conclude that, for a particular specification of backness and height, as many potential vowels are claimed to exist as there are recognized articulatory

[3] It is worth remarking that the purely phonetic evidence—i.e., the articulatory and acoustic evidence as opposed to the phonological data—does not force upon the student the analysis proposed by L, but allows for several alternatives, given our present knowledge. L himself observes that his framework disregards the 'state of the glottis ... which occurs in aspirated sounds' (19—here 'aspirated' is taken in its traditional sense). He mentions in this connection the important X-ray work of Kim 1970, where degree of glottal opening has been shown to correlate with aspiration. But L does not use Kim's facts, on the grounds that 'it is possible to derive a more appropriate set of natural classes for use in phonological descriptions' (19). In the absence of a discussion of Grassmann's Law, this remark falls somewhat short of carrying complete conviction.

places. L recognizes at least six such places as contrasting on the systematic phonemic level (92); hence one expects to find, e.g., six contrasting types of high vowels differing only in having labial, dental, alveolar, post-alveolar, palatal, or velar points of articulation, respectively. Examples of such contrasts among vowels are not documented in the literature; and it surely is no accident that in Table 51, where the proposed dual specification of speech sounds is illustrated, L fails to provide examples of vowels for four of seven feature complexes cited, although his system clearly provides for such vowels. In sum, the proposal that vowels be specified in terms of the articulatory-place feature implies that there are a great many more types of possible vowel sound than anyone has observed. Since L offers no reasons for believing that such sound types are potentially possible in human language, the fact that they have never been found suggests that the claim is false, that vowel sounds are not to be characterized in terms of the articulatory-place feature.[4]

3. Proposals to resolve the difficulty noted above have been made in the past. Ultimately all of these go back to Jakobson 1939, who first suggested that Articulatory place should be dispensed with as a feature in the phonetic framework. Jakobson coupled this with the further proposal that, in the universal feature framework, the features Back, Height, and Articulatory place be replaced by two new features GRAVE and COMPACT. According to his proposal, velar consonants and the low back vowel [a] are grave and compact; palatal consonants and the low front vowel [æ] are acute (non-grave) and compact; dentals and the high front vowel [i] are acute and diffuse (non-compact); labials and the high back vowel [u] are grave and diffuse. It is this proposal that L appears to have in mind when he writes: 'At first sight it might seem appropriate to consider back to be equivalent to velar, and front equivalent to post-alveolar or palatal ... But we do not achieve a satisfactory solution by completely collapsing the backness feature within the articulatory place feature in this way. In the first place, it is plainly wrong to consider low back vowels to be velar sounds. Second, if we do not have separate, additional features for vowels, we cannot consider consonants with secondary articulations to have added vowel-like characteristics' (79–80).

I agree completely with this argument; in fact, it is substantially identical with the argument which Chomsky and I advanced in *Sound pattern of English* (306–8). I expected, therefore, to find at this point in the book an explicit argument against the solution proposed in *SPE*—i.e., to eliminate Articulatory place as a feature and make do with the (slightly redefined) features High, Back, and Low. Although this expectation is not borne out, an argument against the *SPE* solution does appear in L's book. It is found in the last chapter, where he lists his objections to the *SPE* framework and compares it with his own. Since these objections are related to general views that L holds with regard to the nature of vowels, it is useful to begin this discussion by briefly examining his picture of vowel systems.

In his discussion of vowels, L observes that it has long been traditional to characterize them in terms of the highest point of the tongue in a mid-sagittal section of the vocal tract.

[4] In our most recent discussions of the universal framework of features, K. N. Stevens and I have somewhat modified the characterization of vowel height. In effect, we are now inclined to replace TENSE and LOW by two mutually exclusive features, CONSTRICTED PHARYNX and ADVANCED TONGUE ROOT (cf. Perkell 1971). These modifications, however, have no bearing on the issues under discussion here.

Since the tongue can be moved freely in both a vertical and a horizontal direction, it is only natural that phoneticians have described the location of the highest point of the tongue in terms of the two dimensions Height (vertical) and Front–Back (horizontal). But L believes that the traditional definitions of the terms used 'are often not in accord with the facts' (67). Moreover, he points out that one can describe certain vowels 'as differing simply in terms of the single parameter called tongue height only by neglecting large and varied differences in the front–back dimension' (69). Finally, he objects (unjustifiably, I believe) to the fact that when vowels are characterized in terms of the highest point of the tongue, certain vowels that 'form a series of approximately equal auditory steps' are 'far from equidistant' in terms of tongue height. He concludes with the remark: 'Considering all these difficulties, it is difficult to understand how phoneticians could persist in considering that the traditional articulatory categories provide an adequate specification of vowels' (69). Nonetheless, he does not propose to replace the traditional features: 'In view of the complicated relationships between the traditional terms and any of the possible sets of measurements, we might well wonder whether these terms provide the most appropriate basis for phonological features. But there seems to be no doubt not only that linguists do manage to use these labels in a reliable way, but also that language works in terms of them' (74).

L is not impressed with the evidence, offered in *SPE*, for the advantage of replacing the multi-valued tongue-height feature by two binary features High and Low. In fact, he presents the three arguments below for preferring a multi-valued tongue-height feature:

(a) He feels that 'the multi-valued system shows that there is a relation between possible vowel heights of a kind that cannot be stated in binary terms ... there is no way in which a binary notation can [express the fact] that the change from low to mid involves the same process as the change from mid to high. The notion that there is an ordered relationship between vowel heights is a claim that is made by a multi-valued system and not by a binary one. This claim is important in many phonological descriptions of both English (cf. Ladefoged 1967; Foley 1971; Labov 1971) and other languages' (103). As I am unable to include in this review a detailed examination of the three cited works, I can only record my opinion that the arguments adduced in them do not conclusively demonstrate the need for a multi-valued height feature.

(b) L remarks (103) that he does not know how the *SPE* system 'would account for the four front vowels of Danish', which according to him contrast in tongue height. This remark is puzzling to me, since L is obviously aware of the fact that the four Danish front vowels are only a subset of the five front vowels of English, and would therefore be characterized, following *SPE*, by recourse to the TENSENESS feature:

	i	*e*	*ɛ*	*æ*
high	+	−	−	−
low	−	−	−	+
tense	+	+	−	−

For some reason, L does not believe that this is the correct solution, but he presents no facts or arguments of any sort to support this belief. This objection to the *SPE* solution can, therefore, not be counted heavily.

(c) Finally, L claims that 'as far as vowels are concerned, languages work partially in auditory terms ... and partially in physiological terms' (103). I discuss this claim in §4, below. My conclusion there is that the evidence in favor of the claim is weak, and that there is no compelling reason to accept L's proposal over that of *SPE*.

In sum, none of the three arguments advanced against the *SPE* solution appears to establish the need to view vowel heights as a multi-valued feature rather than as the result of the two binary features.[5]

[5] There is a further benefit to be derived from dispensing with the Articulatory-place feature. As shown in Table 59, where L summarizes in convenient fashion the entire proposed feature system, the Articulatory-place feature differs from all others with respect to the number of values it may assume. This feature can assume six values; one other feature

4. L writes: 'It is perfectly possible to describe all the systematic phonetic differences which occur among languages in terms of the sound-producing mechanism. But in some instances this does not seem to be an appropriate way of characterizing the features underlying the contrasts. Correct description of the position of the tongue in vowels is extremely difficult (and not as given in traditional physiological phonetic texts); but differences among vowels are fairly easy to state in acoustic terms. Similarly, some consonants can be grouped together on an acoustic parameter much more easily than in physiological terms. Furthermore, although we could (with difficulty) characterize all possible systematic phonetic contrasts entirely in physiological terms, it would be ridiculous to overlook the fact that some phonological rules work in terms of acoustic properties of sounds ... But it is important to note that we do not have the choice of thinking either in acoustic or in physiological terms. The patterns that arise in the sounds of a language are due to intersecting causes. At least two quite different kinds of features are needed to explain them. Some patterns can be explained in terms of acoustic events, others in terms of articulatory events. Thus, on the one hand, there is no doubt that *p* and *k* go together in the formation of patterns in some languages; this is because of their acoustic similarity, and no amount of guesswork is likely to lead to establishing anything in common in the neural commands to the speech organs which make them. But on the other hand, patterns exhibited in the formation of compounds such as *mp nt ŋk* are obviously due to articulatory constraints; and it is difficult to state rules concerning them in terms of meaningful acoustic features' (4–5)

I have quoted this lengthy passage because I want to make sure that I have not misrepresented L's views on this matter. He claims that some of the contrasts in speech are acoustic in nature, while others are articulatory (physiological); but it is important to realize that he does not deny that each acoustic property of speech has an appropriate articulatory correlate and vice versa. He observes only that, for some features, the acoustic correlate is much simpler to state than the articulatory one; and he believes that this difference in complexity is of such importance that he proposes to subcategorize features into two distinct classes, those with simple articulatory correlates vs. those with simple acoustic correlates. It is, of course, obvious that features can be subcategorized in a myriad of ways; e.g., one might propose to distinguish between features involving only the glottis vs. those that do not; or between features affecting primarily the first formant vs. those affecting the second formant, etc. But such

(Height) can assume four values; but all other features can assume only two or three. Since it is desirable to restrict as much as possible the variety of features admitted into the framework, the elimination of the only six-valued feature is clearly a move in the right direction, especially since it can apparently be made without complicating the framework in any other fashion. It is unfortunate that L does not seem to take abstract, theoretical considerations of this sort into account, for the purpose of theory construction in every science is precisely to limit the number of answers that one might potentially accept in response to a given question. By failing to scrutinize the abstract structure of his theoretical apparatus, L deprives himself of a tool that has been very useful elsewhere, and is quite likely to perform equally well in phonetics.

subcategorization is useless unless it allows us to understand aspects of language which, without it, would remain unexplained. Since such a demonstration is not given in the book, L's claim that 'as far as vowels are concerned, languages work partially in auditory terms ... partially in physiological terms' (103) is somewhat misleading. A more accurate statement might have said that, among the vowel features proposed by L, some have simpler acoustic than articulatory correlates. So stated, this hardly qualifies as a fundamental issue for debate.

In the passage quoted at the beginning of this section, L remarks that certain contrasts can be described in articulatory terms only with great difficulty. It hardly needs saying that the fact that something is difficult to do is no proof of its impossibility: one need only watch a sports event in which records are broken, or a performance of a circus acrobat. But apart from these somewhat abstract, general considerations, L's examples of contrasts that are difficult to characterize in articulatory terms are not particularly persuasive. Thus, in the quoted passage, he refers to the fact that in some languages labial and velar consonants constitute a natural phonological class (as opposed to the dentals and palatals), in spite of the fact that there is nothing 'in common in the neural commands to the speech organs which make them'. But this is also true, mutatis mutandis, of non-nasal sounds (as opposed to the nasals); yet L finds no difficulty in characterizing these in purely articulatory terms. Dentals and palatals can be naturally characterized (as was done in *SPE*) as sounds produced with the active raising of the blade of the tongue toward the roof of the mouth (coronal), in contrast to labials and velars which are produced without participation of the tongue blade (non-coronal). The articulatory definition of the class of non-coronals is, therefore, quite parallel to that of the class of non-nasals. Since there is no problem with the latter, L owes us an explanation as to why he perceives a problem in the former.

Matters are no clearer with regard to the other acoustic feature which L discusses in detail in this connection, that of vowel Height. He explains that the traditional description of vowels in terms of the height of the highest point of the tongue is wrong in the case of the back vowels of Ngwe [u o ɔ a], because while these 'form a series of approximately equal auditory steps ... , the highest points of the tongue are far from equidistant' (69). But L does not justify this requirement on features, nor does he refer to places in the literature where the requirement is justified. I find it plausible that, auditorily, [u] is closer to [o] than to [ɔ]; but I am unable to say whether the distance between [u] and [o] is greater than, smaller than, or the same as that between [o] and [ɔ]. It is, therefore, not obvious to me that one can meaningfully speak of 'auditorily equal steps' among speech sounds. Since L's argument presupposes that this is a meaningful concept, it would have been useful to have some evidence for the psychological reality of the auditory distance measure he required. Since no evidence is cited, L's case rests on nothing more substantial than the reader's willingness to suspend disbelief and to take L's word that there is no serious problem here.

That there is indeed a problem, however, is strongly suggested by the results of investigations by Stevens 1968, who has shown that, in general, variations along a given articulatory dimension (say tongue height) are not related linearly to variations along the corresponding acoustic dimension (i.e. first-formant frequency). In fact, he has shown, with respect to a number of features, that there is a region where small variations in the articulation correspond to large variations in the acoustic correlative. This region is usually intermediate between two other regions where sizeable articulatory variations produce only minor acoustic effects. This picture of the phonetic character of the features has great plausibility: it implies in effect that, in producing speech sounds, the speaker has great latitude as long as he manages to avoid those regions along the articulatory continuum where small variations bring about large changes in the acoustic output. In view of this, there is reason to doubt that the concepts of auditory and articulatory distance can be used in the fashion proposed by L as conditions on features. In any case, the burden of proof here is clearly on L. His

claim that the articulatory specification of tongue height is inadequate is not established; nor can the facts concerning tongue height adduced by him be taken as evidence showing the essentially 'acoustic character' of the feature of tongue height.[6]

Moreover, a deeper issue is involved here. By utilizing a single set of features to characterize both the articulatory and the acoustic properties of speech sounds, as well as the behavior of speech sounds in rules, we are claiming that there is a single principle which explains the regularities to be observed in these three, superficially quite disparate domains. We claim in particular that a special relationship holds between specific pairs of acoustical and articulatory properties of speech sounds (e.g. between first-formant frequency and tongue height). Since this is a very specific claim, it should in principle not be too difficult to contravert it by finding appropriate counter-examples. As a matter of fact, it turns out, on the one hand, that the claim can readily be maintained for large bodies of data drawn from a wide variety of languages (cf., e.g., the data gathered in Jakobson, Fant & Halle 1963); and, on the other hand, that many of the counter-examples adduced are more apparent than real, once they are subjected to careful scrutiny (cf. the discussion in the paragraphs immediately above). Under these circumstances, it does not seem particularly advisable to give up our claim, especially since anyone who gives it up must immediately explain why it appears to hold in so many instances.

There are, of course, many unsolved questions that arise in connection with the claim. It is my guess—and the decision to pursue a particular line of inquiry, rather than another, never rests on anything more solid than a researcher's guess—that progress is to be made by trying to answer these questions, rather than by giving up the claim and thereby declaring the questions uninteresting. Ladefoged currently sees the matter in a different light. The history of the science of language during the next few decades will decide which of these two opposite guesses was more nearly correct.

REFERENCES

CHOMSKY, N., and M. HALLE. 1968. The sound pattern of English. New York: Harper & Row.

COPE, A. T. 1970. Zulu tonal morphology. Journal of African Languages 9.111–52.

HAUDRICOURT, A. G. 1961. Bipartition et tripartition de tons dans quelques langues d'Extrême Orient. Bulletin de La Société de Linguistique de Paris 56.163–80.

JAKOBSON, R. 1939. Observations sur le classement phonologique de consonnes. Proceedings of the 3rd International Congress of Phonetic Sciences, 34–41. Ghent.

——; G. FANT; and M. HALLE. 1963. Preliminaries to speech analysis. 5th printing. Cambridge, Mass.: MIT Press.

KIM, C.-W. 1970. A theory of aspiration. Phonetica 21.107–16.

LISKER, L., and A. ABRAMSON. 1964. A cross-language study of voicing in initial stops: acoustical measurements. Word 20.384–422.

PERKELL, J. 1971. A note on laryngeal features. MIT Research Lab. of Electronics, Quarterly Progress Report, 101.123–9.

STEVENS, K. N. 1968. The quantal nature of speech. Mimeograph, MIT.

[6] It is perhaps worth recalling my argument, above, that L is incorrect in viewing tongue height as a multi-valued feature. If this argument is accepted, then yet another reason for regarding vowel height as an acoustic feature disappears.