WHAT ARE LINGUISTIC SOUNDS MADE OF?

PETER LADEFOGED

University of California, Los Angeles

Linguistic phonetic aspects of languages can be described in terms of about 17 articulatory parameters, and/or a similar number of acoustic parameters. Descriptions of phonological patterns in languages involve features that are not in a one-to-one relationship with these phonetic parameters, and that cannot account for some linguistic phonetic differences among languages. Speakers and listeners producing and interpreting linguistic events probably use something like the proposed phonetic parameters. There is no necessity for most phonological features to be part of mental representations.*

When we give a description of a spoken language, what are the linguistic phonetic parameters? I want to suggest that these are not things like features, but rather things like formant frequencies or parameterized vocal-tract shapes. Despite the conventional wisdom, we cannot be content with specifications of linguistic phenomena in terms of physical scales representing features like those proposed by Chomsky & Halle 1968. Even the considerably better phonological features which I have proposed earlier (Ladefoged 1971, 1975) are far from primitive linguistic phonetic parameters. This point is not made clear in any of these earlier works—and, indeed, as far as I am concerned, was not fully appreciated. I hope to make it clear now that linguistic phonetic descriptions require about 17 articulatory parameters, with about the same number of acoustic parameters. Other aspects of linguistic descriptions, such as accounts of sound patterns within languages, are undoubtedly best stated in terms of phonological features; and if these descriptions are to be explanatory, the features must relate to articulatory or auditory (or cortical) phenomena. But phonological features are certainly not sufficient for specifying the actual sounds of a language; nor are they in a one-to-one relationship with the minimal sets of parameters that are necessary and sufficient for this purpose.

1. We can get a first approximation to a minimal set of articulatory parameters by considering those that have been used in computer programs that synthesize speech. Some years ago, Coker, Umeda & Browman 1973 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.

When we consider a wider range of languages, we must increase the number of parameters. An attempt to list such a set of articulatory parameters is given in List 1. I will not consider all the items on this list: it will be possible to see how the

* Many of the UCLA Phonetics Lab group have hacked critically at drafts of this paper in a series of lab meetings. Their vociferous comments have been a great help. I also received several useful comments from David Isenberg.

This paper represents the Presidential Address given at the 1978 Annual Meeting of the Linguistic Society of America.

1. Front raising	10. Lip width
2. Back raising	11. Lip protrusion
3. Tip raising	12. Velic opening
4. Tip advancing	13. Larynx lowering
5. Pharynx width	14. Glottal aperture
6. Tongue bunching	15. Phonation tension
7. Tongue narrowing	16. Glottal length
8. Tongue hollowing	17. Lung volume decrement
9. Lip height	

LIST 1.

parametric approach differs from more traditional linguistic descriptions by considering only those parameters that specify the position of the tongue. Furthermore, the set of parameters listed is only a first approximation to those required. We do not yet know enough to be able to specify all and only the parameters required for linguistic contrasts; but I hope this list will be sufficient to give some impression of a possible set of phonetic parameters, and to show their relationship to more familiar phonological features.

The first two parameters, front raising and back raising, are illustrated in Figure 1. They specify the position of the body of the tongue. The front-raising parameter

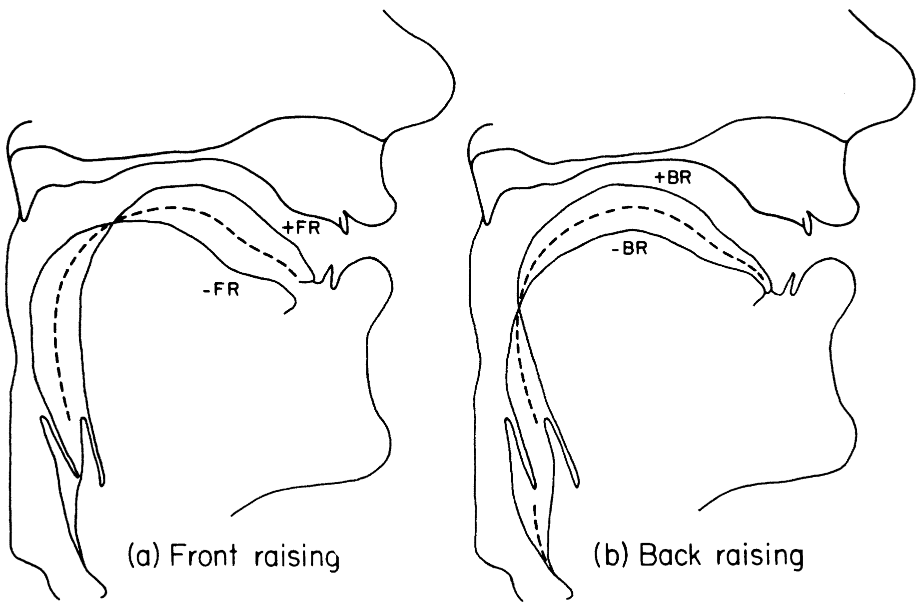

FIGURE 1. Positive and negative values of (a) the front-raising parameter, and (b) the back-raising parameter. The reference position is indicated by a dashed line.

may be thought of as a movement from something like the position occurring in [o] to something like that in [i], as shown in Figure 1a. The back-raising parameter specifies a movement from approximately [ɑ] to [u], as shown in Figure 1b. In each case, the movement should really be thought of as a deviation from the reference position of the tongue, so that terms like 'front raising–lowering' and 'back raising–lowering' might be more appropriate. Both parameters have been defined in formal terms (Harshman, Ladefoged & Goldstein 1977) as deviations (in centimeters) of the tongue position of an average speaker from the position of that speaker's tongue in a reference position.

The tongue positions of all the non-rhotacized vowels of American English may be specified pretty accurately in terms of these two parameters. For example, Figure 2 shows how the vowel [u] may be thought of as a certain amount of back

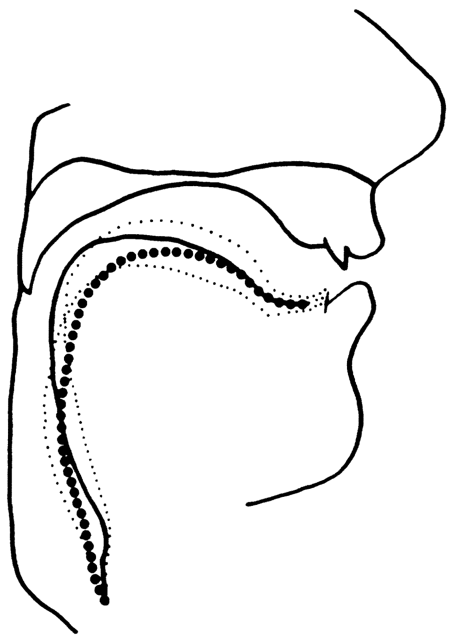

FIGURE 2. Reconstruction of the vowel /u/ as in *who* (solid line). The heavy dotted line indicates a reference position for the tongue, and the two light dotted lines indicate the deviations from the reference line of the two parameters that sum to give the deviation for /u/. The two dotted lines cross near the uvula. (For reasons of clarity, the epiglottis is not shown.)

raising, combined with a negative amount of front raising (i.e. a deviation below the reference line) which keeps the front of the tongue down and moves the body of the tongue further back. When the deviations corresponding to these parameters are added together, the position of the tongue for [u] results.

Recent work (partly reported in Ladefoged et al. 1978) has indicated that these two parameters can be used to describe the tongue shapes of vowels in other languages reasonably well. Additional parameters that will be discussed later are necessary to account for some distinctive shapes; but the two parameters shown (or something very like them) will probably account for more of the variance found in the vowels of the languages of the world than any other two parameters for specifying tongue shapes.

This claim instantly invites comparison with other systems of describing vowels, such as a more traditional description in terms of the position of the highest point of the tongue. The trouble with the traditional system is that it defines the location of only one point on the tongue—and there is no algorithm for describing the position of the rest of the tongue, given just that information. Figure 3 shows two

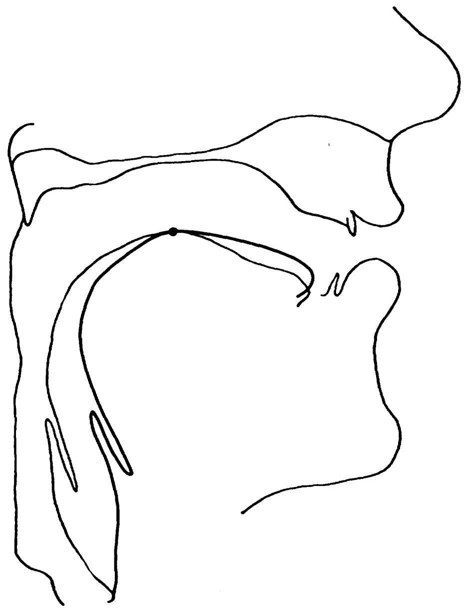

FIGURE 3.

possible tongue positions that have the same highest point: given only the location of the solid point, there is no way of determining which of these (or many intermediate) shapes is being described. It may eventually be possible to use the highest point of the tongue to refer to unique, actually observed, tongue positions. But no one has demonstrated a method for doing this, and it is impossible to say how much of the variance among vowels could be thus accounted for.

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. It seems, in fact, as if they 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–lowering 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 these parameters (and perhaps the others that I will be discussing) describe higher-level cortical control functions. That is, we may think of them as the underlying parameters that deter-

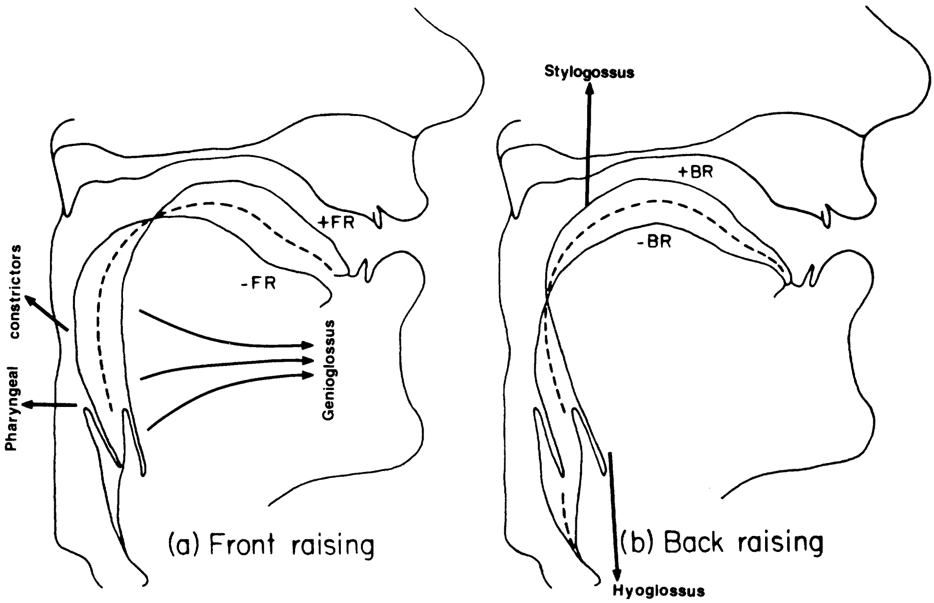

FIGURE 4. The principal muscular forces involved in (a) front raising and (b) back raising.

mine the synergistic actions which are required for the skilled motor movements that occur in speech.

But, we might well ask as linguists, does any of this have any explanatory power from our point of view? What is important to us is whether these parameters help us account for phonological phenomena. This may be considered by seeing how they divide vowels into classes. Figure 5 (based on data in Ladefoged et al. 1978)

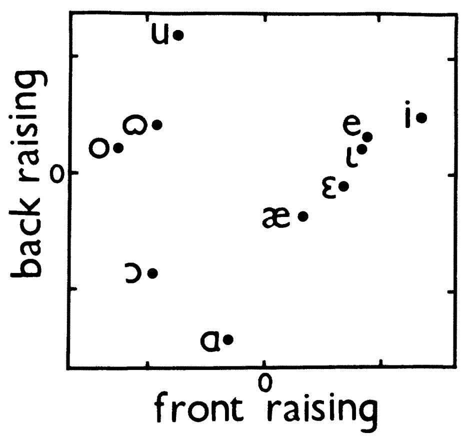

FIGURE 5.

shows the degree of front raising and back raising in ten American English vowels (mean values for five speakers). There is a very general similarity between the arrangement of the vowels in this figure and their location in a traditional vowel chart. The front-raising parameter clearly separates front vowels from back vowels.

But the back-raising component, considered as a single physical scale, is not very useful in explaining observed vowel patterns, or in writing phonological rules for alternations of vowels (although it does help explain articulatory similarities such as that between low back vowels and pharyngeals).

I have discussed these two articulatory parameters at some length in order to provide a good example of the lack of a match between linguistic phonetic descriptions and traditional phonological units. It is worth remembering that these articulatory parameters have been set up simply to account for linguistic differences among utterances, and in this sense are linguistic primes. This is a point to which we will return when we have discussed, somewhat more briefly, some of the other parameters in List 1.

The next two parameters listed, tip raising and tip advancing, have a more straightforward function. They, too, can be defined in quantitative terms as deviations (in centimeters) from a reference position for an average speaker. Figure 6 shows the two-dimensional movements of the tip of the tongue associated

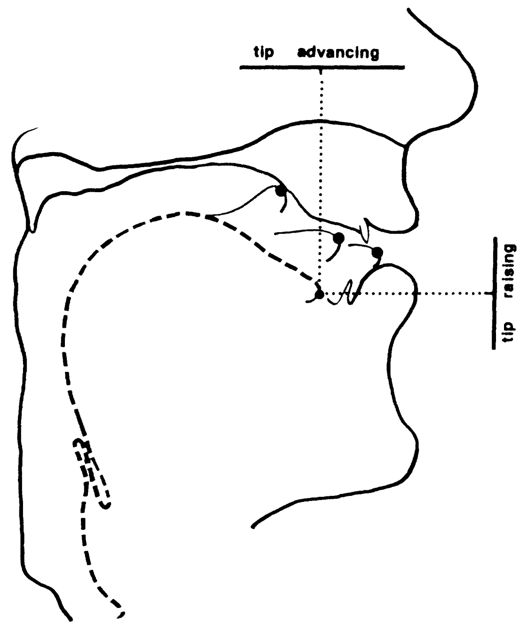

FIGURE 6.

with retroflex, alveolar, and interdental positions. As a first approximation, these movements may be considered as specifying variations in the position of the tongue that are independent of those specified by the parameters for the body of the tongue. Perhaps, in years to come, when we may have a larger body of good phonetic data, we will be able to take into account the correlation between movements of the tip of the tongue and those of the rest of the tongue. This correlation is particularly obvious in gestures such as sticking the tip of the tongue as far as possible out of the mouth—a maneuver doctors request when they want the root of the tongue pulled forward out of the way, so that they can view the larynx. But even in the comparatively small movements involved in speech, there may be interactions. As Ohala 1974 has observed, alveolar and dental consonants sometimes seem to cause a lowering of the back of the tongue.

The phonological correlates of movements of the tip of the tongue are readily apparent. But it is worth noting that features like Coronal (or Alveolar) can be defined only in terms of both tip raising and tip fronting. Given this particular set of articulatory parameters, there is no way that Coronal (or Alveolar) can be interpreted in terms of a single physical scale. Of course, it would be possible to say that linguists must include in their phonetic descriptions some additional, ad-hoc parameter, such as the distance to which the blade of the tongue is raised from its reference position. But this would be pointless, because this parameter would be fully predictable from those already needed. The parameters being described constitute a necessary and sufficient set to account for all linguistic differences between utterances.

Let me emphasize that I am not suggesting that terms like Coronal or Alveolar should be replaced in phonological descriptions by terms like Tip raising and Tip advancing. When describing the sound patterns of languages, we will want to refer to natural classes defined in terms of conventional phonological features; and these features must refer to observable phonetic phenomena. But I do advocate that, when we make a phonetic description of a language, we should not do so by trying to interpret each feature in terms of a single physical scale. We must be able to describe sound patterns in terms of phonological features; but we must also be able to map these features onto the basic linguistic phonetic parameters. The form of the mapping rules involved will be discussed later in this paper.

Similar points can be made with respect to other phonological features that are sometimes used for distinguishing alveolar dental sounds. The two degrees of freedom required for tongue-tip movement can be combined with the parameters required for the specification of the body of the tongue, to distinguish between apical and laminal sounds (or [$\pm$distributed] if that terminology is preferred). The apical and laminal categories are abstractions involving more than one of the physically definable parameters; they are not themselves part of the set of minimal phonetic parameters.

Additional parameters are required for specifying other aspects of tongue position. Among these are the variations in pharynx width (tongue-root advancement) that occur in vowels—and perhaps in obstruents—as well as the bunching of the tongue that occurs in /r/ sounds, sometimes with and sometimes without a movement of the tip of the tongue. Figure 7 shows my current best estimate of the deviations in tongue position that can be associated with each of these parameters. The diagram for tongue bunching should be regarded as especially tentative; it is based on an analysis of a very limited number of American English speakers saying words such as *heard*. But these data (together with the observations of Uldall 1958 and Delattre & Freeman 1968), plus our own experience in synthesizing /r/ sounds from an articulatory model, all indicate that the shapes of the tongue occurring in these sounds cannot be produced without a contraction in the pharynx something like that shown.

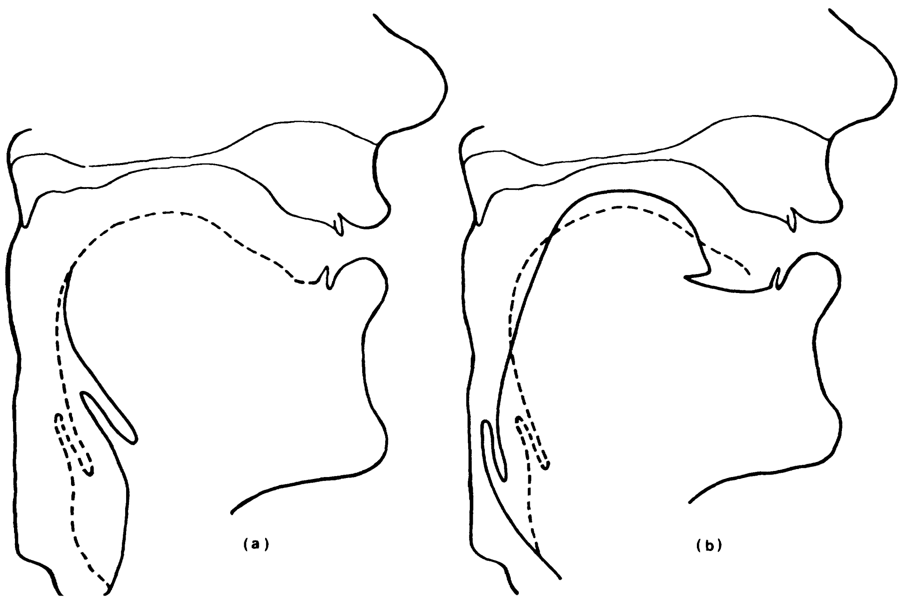

FIGURE 7. The movements of the tongue associated with (a) pharynx width, and (b) tongue bunching.

Both pharynx width and tongue bunching do, in fact, correlate in a fairly simple way with phonological features. Because of lack of data, I cannot say much about tongue bunching, except that it seems to be in a monotonic relationship with the phonological feature Rhotacization (Ladefoged 1975). Pharynx width is worthy of further comment, in that it correlates very highly with the feature Expanded as recently discussed by Lindau 1978, and also offers an interesting insight into a problem that has troubled phonologists for some time. It has never been clear how one should give good phonetic definitions of overlapping phenomena; thus, in languages with pairs of vowels like [i ɩ] and [e ɛ], in which the distinction is said to be relative advancement of the tongue root, one always wants to know: relative to what? Using the parametric approach outlined here, one can give an answer that will produce physical specifications for an average (or any other) speaker. The major aspects of the tongue position are the result of adding deviations from the reference position associated with the front-raising and back-raising parameters, and the variations in the position of the root of the tongue are additional deviations associated with the pharynx-width parameter. Observed tongue shapes are the result of summing the actions of these three (and other) underlying parameters.

On tongue narrowing, the parameter associated with laterals, I have little to say except to express a hope that it can be defined so as to account for dental, alveolar, retroflex, palatal, and velar laterals. This may not be possible; but it is only when we can give a formal account of what part of the tongue is narrowed that we will be able to give a really meaningful definition of the physical correlates of the feature Lateral.

Tongue hollowing is included here as an additional parameter to account for other variations in tongue shape in the coronal plane (the view from the front, as opposed to the more traditional sagittal view of the vocal organs). There are clear variations in the degree of hollowing and doming of the tongue, e.g. in English [s] and [ʃ]. But I cannot give a quantitative description of these phenomena. Again let me emphasize that the parameters listed are illustrative of my approach, rather than definitive of what phoneticians have observed. The first two parameters were algorithmically derived from a limited set of English vowels (Harshman, Ladefoged & Goldstein); but the remainder are simply my best estimates. The available evidence, however, suggests that these parameters are far more likely to be what a speaker actually controls in producing sentences than any of the current phonological features.

Many of the remaining parameters in List 1 could have been used to demonstrate the relationship advanced here between phonological and phonetic units. But by now it should be clear that the necessary and sufficient set of articulatory parameters required for characterizing linguistic contrasts is not identical with the set of features required for characterizing phonological patterns. A similar point can be made by reference to the acoustic parameters of speech.

2. Authorities differ on the precise composition of the minimum set of acoustic parameters required for synthesizing human speech. The set of parameters used by the OVE 3 Speech Synthesizer (Liljencrants 1968) is shown in List 2. The adequacy of something like these parameters for describing speech was demonstrated by Fant and his colleagues about 15 years ago.

1. Voice source frequency
2. Voice source amplitude
3. Frequency of formant one
4. Frequency of formant two
5. Frequency of formant three
6. Bandwidth of formant one
7. Bandwidth of formant two
8. Bandwidth of formant three
9. Amplitude of nasal formant
10. Frequency of nasal formant
11. Amplitude of aspiration
12. Amplitude of fricative source
13. Frequency of lower fricative pole
14. Frequency of upper fricative pole
15. Relative amplitude of fricative poles

LIST 2.

It might be objected that all the languages synthesized by Fant and his associates belonged to the Indo-European family, and we do not know if these parameters are sufficient to specify the complete range of possible linguistic contrasts. Personally, I feel that the consonant combinations that occur in Russian are as tough a test for acoustic specification as anything found elsewhere. But in order to demonstrate that even so-called exotic sounds can be adequately specified in this way, I have synthesized two non-Indo-European phrases (Ladefoged, MS). The first is part of a greeting exchange in Yoruba and contains a labial velar stop (cf. Ladefoged 1968): [ɛkârɔ̀o ó a duk͡pɛ́] 'Good morning. Thank you.' The second is Zulu, and contains two clicks: [gʗìba isiɔ^{h}úmo] 'I saw an antelope.' There are some minor imperfections in the specification of these phrases, but nothing indicates that any of these sounds cannot be specified in terms of appropriate time-varying values of these or some similar set of acoustic parameters.

Some phonological features can be readily interpreted in terms of a single acoustic parameter; thus vowel height is inversely proportional to the frequency of the first formant. But the majority of phonological features are in a many-to-many relationship with the minimal set of acoustic parameters, just as they are with the minimal set of articulatory parameters.

It is always possible to convert a specification of a particular speech sound in articulatory terms into a unique specification of that sound in acoustic terms. To a great extent, it is also possible to reverse this process and convert an acoustic specification into a unique articulatory specification. In a few cases, more than one articulatory configuration will produce a given sound. Atal et al. 1978 have shown that it is theoretically possible to produce some vowels with a variety of different vocal-tract shapes. More importantly to us as linguists, Riordan 1977 has indicated how compensatory articulations may be involved in linguistic change: she found that, in certain experimental circumstances, subjects will attempt to maintain a given vowel quality by using adjustments of the larynx position to compensate for different lip positions. We do not yet know to what extent these compensations are fundamental properties of language.

Work at UCLA (Ladefoged 1979) indicates that possible compensatory articulations are strictly limited. I cannot, for example, find any possible ways in which two distinct tongue positions will produce the same sound without changes also occurring in the positions of the lips or the larynx. It is a matter of common experience that speakers generally use very similar articulations to produce a given sound: they close their lips to produce /p b m/, raise their tongue to the alveolar ridge to produce /t d n/, and so on. In fact, probably the only consonant in American English that does not have a unique articulatory specification corresponding to a particular acoustic specification is /r/, which may be produced with the tip of the tongue up or down, and with variation in the degree of bunching of the tongue. Similarly, with vowels, 'Whenever a speaker produces the vowel /i/ as in *heed*, the body of the tongue is always raised up towards the hard palate. Whenever anyone produces the vowel /ɑ/ as in *father*, the tongue is always low and somewhat retracted' (Ladefoged et al. 1978).

This interconvertibility of articulatory and acoustic descriptions has recently been exploited by members of the UCLA Phonetics Lab (Ladefoged & Lindau 1978). Our work on going from speech sounds back to vocal-tract shapes has progressed to the stage where we can take a short sentence, consisting of predominantly vocalic sounds, and reconstruct a set of vocal-tract shapes that might have produced this utterance. We cannot as yet reconstruct plausible vocal-tract shapes corresponding to true consonants, simply because the appropriate algorithms have not yet been written. At the moment we can handle only such everyday utterances as *We owe you a yoyo*; *How will you woo her away?*; and *We will weigh you.* But there are no theoretical difficulties in going much further.

Since it is always possible to convert an articulatory description into an acoustic one, and vice versa (though not necessarily uniquely), it might appear that either the acoustic parametric description or the articulatory description is redundant. We could say that one or the other comprises the minimal set of linguistic phonetic parameters, but not both—since either set would allow us to make descriptively adequate statements. The articulatory parameters would also serve as a basis for explanatory statements by physiologists concerned with motor movements, and the acoustic parameters as a basis for explanatory statements by those interested in audition. But as linguists, we will want to refer to both sets of parameters. Languages get to be the way they are because of the interplay between articulatory and acoustic (and other) factors. As was noted when discussing the two sets, some phonological features correlate in a simple way with parameters from one set, and others with the other. I used to think (Ladefoged 1971, 1975) that all but a few features (which I called cover features) could be defined in terms of either simple articulatory scales or simple acoustic scales. It now appears to me that this is an oversimplification, and that very few features can be directly correlated with any of the minimal phonetic parameters.

3. Having considered the mapping of articulatory parameters onto acoustic parameters, and vice versa, we must now discuss the way in which systematic phonetic descriptions can be mapped onto parametric descriptions of either kind. As I have been emphasizing, I agree with the standard view that phonologies should describe sound patterns by means of rules linking underlying forms with systematic phonetic descriptions. But the values assigned to the features at the systematic phonetic level are not full descriptions of the sounds. Taken as a set, they are neither necessary nor sufficient to specify what it is that makes English sound like English rather than German. To map features onto articulatory or acoustic parameters, something like a speech-synthesis-by-rule program is needed, to provide additional information. Thus the rules for mapping the three segments (each considered as a set of feature values) in [kʰæʔt] 'cat' have the general form

$$P_i = \alpha f([\text{k}^{\text{h}}]) + \beta f([\text{æ}]) + \gamma f([\text{ʔt}])$$

where P_i is the value for parameter i; $f([\text{k}^{\text{h}}])$ is a function of the feature values of [kʰ] (the particular allophone that occurs in *cat*); $f([\text{æ}])$ is a function of the feature values in the allophone [æ]; and $f([\text{ʔt}])$ of those in that allophone. The variables α, β, γ are time-varying weighting functions corresponding to the degree of coarticulation that occurs in these circumstances. The functions for the allophones may be thought of as fairly straightforward look-up tables. Thus, to specify the position of the body of the tongue associated with [kʰ], we may write:

(1) $\begin{bmatrix} +\text{velar} \\ +\text{stop} \end{bmatrix} \rightarrow P_{\text{front raising}} \; -1.0$

(2) $\begin{bmatrix} +\text{velar} \\ +\text{stop} \end{bmatrix} \rightarrow P_{\text{back raising}} \; +3.0$

Note that we cannot interpret [+velar] by itself, or [+stop] by itself; the front-raising and back-raising parameters must be determined by considering both these features together. Note also that these mapping rules have no psychological reality. They are simply ways of relating one set of linguistic facts (phonological descriptions of the sound patterns) to another (phonetic descriptions of the sounds of one language as opposed to another).

Certain consequences of the separation of phonological and phonetic descriptions must be faced. First, it leaves us with fewer guiding principles as to what defining limits can be set on phonological features. If there is no one-to-one correspondence between phonological features and the parameters which a speaker actually uses in producing sentences, then phonological rules may be stated in terms of much more abstract units. This is entirely satisfactory if the object of a phonology is merely to describe the sound patterns observable in a language; these patterns are not necessarily used in any way by speakers of the languages (although many have assumed that they are).

Second, if we go on using the linguistically well-known feature sets which have been found very useful in phonological descriptions, we must do so with the realization that these feature sets—mine, Chomsky & Halle's, or anyone else's—have in no way been proved to be the mental representations used by people when speaking or listening to any language. Most of them are completely unnecessary for adequate descriptions of the behavior of speakers and listeners. 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 representations of what a speaker knows about the social institution called language. They are abstract constructs that can be evaluated only in terms of criteria such as the degree of simplicity and elegance that they permit in descriptions of the data.

A language (i.e., the social institution that permits formalized communication between individuals) may well be most appropriately described in terms of segments and features. When linguists write phonologies, they are (whether they admit it or not) describing properties of social institutions, and only guessing at the mental representations of speakers or listeners. The properties of languages are as they are because of the conflicting activities of different groups: speakers who often want to produce rapid, distorted, articulations—as opposed to listeners, who usually prefer clear, maximally distinguished sounds—and language learners, who are most successful when the patterns in the language can be easily observed and stored. The intersection of these different activities produces a system that can be described in terms of segments and features. The very rich sets of patterns observable in most languages have arisen because of the criss-crossing effects of the systems of speech production, perception, and acquisition. The properties of the abstract thing that we call language are different from those required in a set of rules for sentence production, or from the different set of rules required for sentence perception.

We all have some knowledge of the patterns that occur in the social institution we use for communicating. But it is not clear to me whether this knowledge is necessarily part of what we want to call linguistic competence. Most people (and certainly all linguists) know far more than is required for producing or understanding sentences. We can play language games, make rhymes, and even (after a lot of instruction) read and write; but this competence is not, in general, relevant to the processes of talking and listening. For these activities, we need to operate principally in terms of parameters such as those in Lists 1–2.

4. Returning to the discussion of these phonetic parameters, we should consider the evidence that either set listed may be a minimal set of LINGUISTIC phonetic units. I have simply asserted that these parameters are necessary and sufficient for phonetically characterizingl inguistic contrasts. That they are necessary is evident from attempts to synthesize speech; if we omit the values for one of them, we will be unable to produce certain contrasts. That they are sufficient is a harder claim to justify; but it can easily be disproved by finding counter-evidence. At the moment, there is no reason to believe that such evidence is likely to become available. All the recently described phonological contrasts seem to be new combinations of previously known possibilities rather than totally new phenomena. Thus the contrasts of dental, alveolar, and velar laterals in Melpa and Mid-Waghi (Eastern Highlands, Papua New Guinea), exemplified in 3–4 (Ladefoged, Cochran & Disner 1977), involve no new parameters:

(3) Melpa

	MEDIAL		FINAL	
Dental	*kial̪t̪im*	'fingernail'	*wal̪*	'knitted bag'
Alveolar	*lola*	'to speak improperly'	*bal̪*	'apron'
Velar	*paǥa*	'fence'	*raǥ̥*	'two'

(4) Mid-Waghi

al̪aal̪a 'again and again'
alala 'to speak improperly'
aǥaǥe 'dizzy'

Similarly, the contrasting bilabial trills in Kele and Titan (Austronesian, Papua New Guinea), exemplified in 5–6, can be described in terms of the lip and vocal-cord parameters in List 1, or the formant and larynx-source parameters in List 2:

(5) Kele

Bilabial	*mBin*	'vagina'	*mBulin*	'your face'	*mBɛŋkei*	(a fruit)
Lingual	*nruwin*	'bone'	*nrilɛŋ*	'song'	*nrikei*	'leg'

(6) Titan

Bilabial	*mBulei*	'rat'	*mButukei*	'wooden plate'
Lingual	*ndruli*	'sandpiper'	*ndrakeʔin*	'girls'

The different voice qualities in Mpi (Harris, p.c.), illustrated in Table 1, can be accounted for in terms of the glottal parameters shown in List 1. In this case, the

	PLAIN	LARYNGEALIZED
1.	˩˧ *si* 'to be putrid'	˩˧ *sḭ* 'to be dried up'
2.	˩ *si* 'blood'	˩ *sḭ* 'seven'
3.	˥˧ *si* 'to roll'	˥˧ *sḭ* 'to smoke'
4.	˧ *si* (a color)	˧ *sḭ* (classifier)
5.	˦˥˦ *si* 'to die'	˦˥˦ *sḭ* (man's name)
6.	˥ *si* 'four'	˥ *sḭ* (man's name)

TABLE 1. Contrasts between the six tones and between the plain and laryngealized vowels in Mpi (Southern Lolo branch, Tibeto-Burman, Northern Thailand).

acoustic parameters shown in List 2 may have to be expanded, to allow for differences in glottal pulse shape that are apparently controllable in Mpi. But Laver 1977 has shown that a very remarkable range of phonation types can be synthesized in terms of comparatively simple combinations of acoustic parameters.

Potentially, a far larger range of speech sounds could be contrastive. Several other phonological contrasts might occur in as yet undescribed languages, and some of these might be very difficult to describe in terms of the present parameters. For example, no known languages use lateral movements of the tongue between the lips, but any child can make a noise of this type. To the best of my knowledge, no language uses buccal fricatives in which movements of the cheeks produce an egressive airstream, though these are also fairly simple to learn. All such sounds might be considered part of the 'phonetic capabilities of man' (Chomsky & Halle). It would be completely improper to disregard them simply because they have not yet been observed in some language. However, until I am proved wrong by events, I will contend that they need not be characterizable in terms of either set of parameters, on the grounds that they are too hard to integrate into a spoken language.

I am, of course, aware that this is a vague pronouncement, a sort of hand-waving to indicate that something might be done to make the theory being proposed more testable. In reality, the validity of these parameters (as in much of linguistics) is not a scientific notion that is dependent on an empirically testable hypothesis. As Abercrombie 1956 has pointed out, tests that involve knowing all present, past, and future languages are obviously pseudo-procedures. A theory concerning the phonetic capabilities of man is, inevitably, just a description of the known data on the basis of which one can provide speculative (but never scientifically proved) explanations that predict what is likely to be observed in the future. Linguists, like any other group of scientists/observers, should seek explanations for the regularities they observe; but they should not be worried if their explanations are merely predictive of future observations and not otherwise testable. However, this view does not absolve us from the responsibility of expressing our observations, whenever possible, in terms of numbers that can be shown to be valid, reliable, and significant. Linguistic phonetic descriptions can do this appropriately by reference to the parameters in Lists 1–2.

5. Having spent much of this paper discussing contrasts within individual languages, I must now consider how to describe measurable phonetic differences between languages. The sounds of one language may differ from those of another because of the phonetic value of the segments; these differences are as much linguistic properties of the languages as are those in the sound patterns, often more fully described. As linguists, we tend to get so involved with describing the phonology of, say, English or Danish, that we forget to point out that many of the sounds of English are not the same as the similarly specified sounds of Danish. Thus even the monumental work of Chomsky & Halle is only a description of the patterns, not of the sounds. The theory may make it possible to give precise descriptions of the sounds of English. But the fact remains that *SPE* does not tell us all we need to know about the phonetic properties of a vowel that is specified as, e.g., [+high, −low, −back]. We cannot tell if it sounds the same as a vowel that may be similarly specified in a description of Danish. As phoneticians have long known, /i/ in English is not the same as /i/ in Danish, and a complete linguistic description of each language must make this evident.

The inadequacy of current phonological theories becomes more apparent when we consider sounds like the velar ejectives in Hausa and Navajo. These consonants may be given the same label and written with the same symbol [k'] in a phonetic transcription. But they do not sound the same. If a Navajo speaker used a Hausa velar ejective while speaking Navajo, it would sound as if he had a foreign accent. It is very difficult to describe differences of this kind in terms of phonological features; but if there is a noticeable difference between two sounds in different languages, such that either of them would sound foreign if it were used in the other language, then this difference is part of the linguistic facts of each language.

I will now consider two cases in which there are measurable phonetic differences between languages that should be evident from full descriptions of each language. In neither case can the differences between languages be taken into account by some notion of variation in the 'basis of articulation'; in fact, this whole concept invariably seems inadequate for discussing differences between languages. I know of no quantified differences between languages that can be handled thus. In every case, when giving a precise account of what makes a particular language sound the way it does, it is necessary to describe the phonetic properties of individual segments.

The first set of data I will use to illustrate this point comes from Disner 1978. She has compared the vowels of Germanic languages, and has been able to substantiate traditional phoneticians' auditory judgments of the phonetic differences among these languages. For example, her plot of the formant frequencies of some of the long vowels of Danish is given in Figure 8. Each ellipse is centered at the

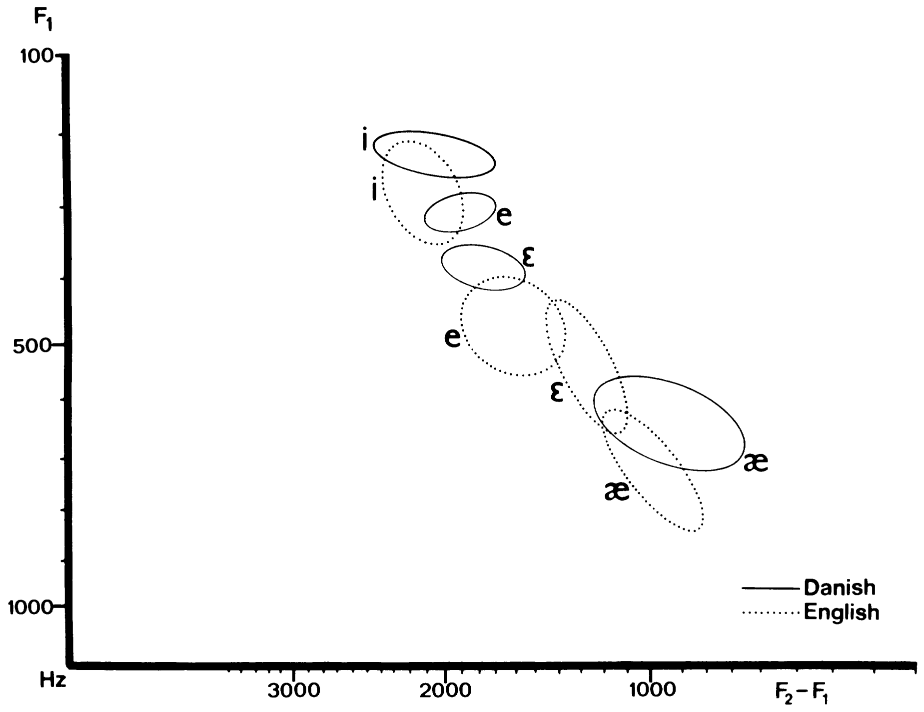

FIGURE 8.

mean of the reported formant frequencies for the vowel (Fischer-Jørgensen 1972) and has axes with lengths of two standard deviations. For comparison, the locations of four of the vowels of English are shown by dotted lines. The frequency values have been plotted on scales such that distances between points reflect perceptual distances. It is obvious that Danish /i/ is higher than English /i/; moreover, the four front unrounded vowels of Danish are unevenly spaced, three of them being much higher than their English counterparts. We cannot say that Danish has a higher basis of articulation than English, because there is no uniformity to the difference between the two languages. Each Danish vowel is higher than its English counterpart, but the difference varies for each pair. We need specific descriptions of each vowel in each language, in order to show how the vowels of one language are phonetically distinct from similarly specified vowels in the other.

In this particular case, the phonetic differences between the two languages can be expressed in terms of scalar values of features such as Height (or High and Low) that have simple acoustic correlates. But the phonological features that have been used to describe sound patterns within languages are, in many cases, not sufficient to account for linguistically significant differences between languages. Thus both Kalaḅari and Hausa, languages of Nigeria, have voiced glottalized bilabial and alveolar stops—sounds usually transcribed as /ɓ ɗ/. Figure 9 shows spectrograms of words containing these sounds preceded and followed by low vowels, and illustrates a considerable difference between the two languages. In the Hausa words, in the upper row, the preceding vowel is marked by irregular vibrations of the vocal cords, and there is at best laryngealized voicing throughout the closure. But in the Kalaḅari words, in the lower row, the implosive sounds are fully voiced throughout the closure, and there is no tendency toward creaky voice or laryngealization.

I have investigated recordings of a number of speakers of each of these languages; there is no doubt that this is a reliable, quantifiable, significant difference between the languages. This difference can be described in terms of the parameters listed in Lists 1–2; but it cannot be handled in terms of the features suggested by Chomsky & Halle. It might be possible to use the features suggested by Halle & Stevens 1971, which involve laryngeal parameters very similar to those in List 1. But those features, like those in List 1, do not enable us to categorize sounds into phonologically appropriate natural classes; e.g., English sounds that differ in voicing must be distinguished in terms of two separate features, which Halle & Stevens call Stiff and Constricted. Again we see that the features that are necessary and sufficient for describing the phonetic properties of languages are not in a one-to-one relation with the features required for phonological descriptions.

The difference between Kalaḅari and Hausa implosives is not known to be contrastive within any one language. But there seems to be no principled reason why it should not be; it is perfectly audible, even to speakers of languages that lack all these sounds. The same is true of several other differences between languages; thus there are reported differences in the kind of lip-rounding that occurs in French and German. The sibilants of Swedish and Polish may also differ; and the tap *r*-sounds in Hausa and Malayalam may involve different degrees of lowering of the

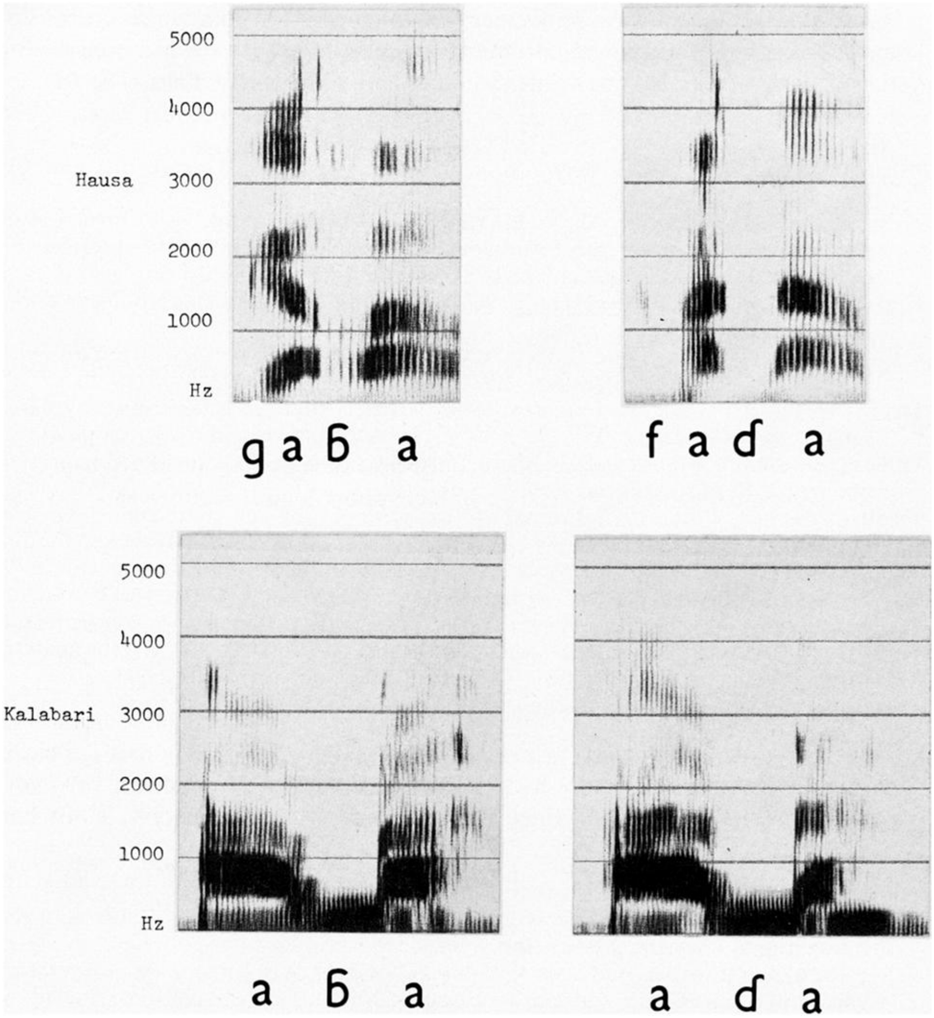

FIGURE 9. Voiced bilabial and alveolar implosives in Hausa /gáɓàa/ 'joint', /fáɗàa/ 'quarrel'; and Kalaḅari /aɓa/ 'to kill her', /aɗa/ 'her father'.

frequency of the third formant. There is no doubt that speakers can make, and listeners can hear, at least some of these differences with complete reliability. Therefore this degree of phonetic detail must be included in linguistic phonetic descriptions of languages.

6. In summary, I have tried to show that the fundamental linguistic phonetic constraints are sets of articulatory or 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 phonological patterns in languages 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 levels, 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 parameters in Lists 1–2.

REFERENCES

ABERCROMBIE, DAVID. 1956. Pseudo-procedures in linguistics. Lecture given at the University of Edinburgh.

ATAL, BISHNU; J. J. CHANG; M. V. MATHEWS; and J. W. TUKEY. 1978. Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. Journal of the Acoustical Society of America 63.1535–56.

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

COKER, CECIL; N. UMEDA; and C. BROWMAN. 1973. Automatic synthesis from ordinary English text. IEEE Transactions, Electroacoustics, AU-21.293–8.

DELATTRE, PIERRE, and D. C. FREEMAN. 1968. A dialect study of American *r*'s by x-ray. Linguistics 44.29–68.

DISNER, SANDRA. 1978. Vowels in Germanic languages. (UCLA working papers in phonetics, 40.) Los Angeles.

FISCHER-JØRGENSEN, ELI. 1972. Formant frequencies of long and short Danish vowels. Studies for Einar Haugen, ed. by E. S. Firchow et al., 189–213. The Hague: Mouton.

HALLE, MORRIS, and K. N. STEVENS. 1971. A note on laryngeal features. Quarterly Progress Report, MIT Research Laboratory of Electronics, 101.198–213.

HARDCASTLE, W. J. 1976. Physiology of speech production. London: Academic Press.

HARSHMAN, RICHARD; PETER LADEFOGED; and LOUIS GOLDSTEIN. 1977. Factor analysis of tongue shapes. Journal of the Acoustical Society of America 62.693–707.

LADEFOGED, PETER. 1968. A phonetic study of West African languages. Revised edition. Cambridge: University Press.

——. 1971. Preliminaries to linguistic phonetics. Chicago: University of Chicago Press.

——. 1975. A course in phonetics. New York: Harcourt Brace Jovanovich.

——. 1979. Articulatory parameters. Proceedings of the 9th International Congress of Phonetic Sciences, Copenhagen, 41–7.

——. MS. The phonetic specification of languages of the world. To appear, Proceedings of the 8th International Congress of Phonetic Sciences, Leeds.

——; ANNE COCHRAN; and SANDRA DISNER. 1977. Laterals and trills. Journal of the International Phonetic Association 7.46–54.

——; RICHARD HARSHMAN; LOUIS GOLDSTEIN; and LLOYD RICE. 1978. Generating vocal tract shapes from formant frequencies. Journal of the Acoustical Society of America 64.1027–35.

——, and MONA LINDAU. 1978. Prediction of vocal tract shapes in utterances. Journal of the Acoustical Society of America 64.S41.

LAVER, JOHN. 1977. Voice quality. Paper presented to the 1st International Phonetic Sciences Congress, Miami Beach.

LIEBERMAN, PHILIP. 1977. Speech physiology and acoustic phonetics: An introduction. New York: Macmillan.

LILJENCRANTS, JOHAN. 1968. The OVE III speech synthesizer. IEEE Transactions, Electroacoustics, AU-16.137–40.

LINDAU, MONA. 1978. Vowel features. Lg. 54.541–63.

OHALA, JOHN. 1974. Phonetic explanation in phonology. Papers from the Parasession on Natural Phonology, 251–75. Chicago: CLS.

RIORDAN, CAROL. 1977. Control of vocal-tract length in speech. Journal of the Acoustical Society of America 62.998–1002.

ULDALL, ELIZABETH T. 1958. American 'molar' *r* and 'flapped' *r*. Revista do Laboratorio de Fonetica Experimental, Coimbra, 4.103–6.

[Received 19 June 1979.]