94

Some Reflections on the Theoretical Bases of Phonetics

M. HALLE AND K. N. STEVENS

Massachusetts Institute of Technology, Cambridge, Mass, USA

"We are concerned with the eternal antithesis between the two inseparable components of knowledge, the empirical and the rational . . . The structure of the system is the work of reason; the empirical contents and their mutual relations must find their representation in the conclusions of the theory. In the possibility of such a representation lie the sole value and justification of the whole system, and especially the concepts and fundamental principles which underlie it. Apart from that these latter are free inventions of the human intellect, which cannot be justified either by the nature of the intellect or in any other fashion a priori." (A. Einstein "On the Method of Theoretical Physics" Spencer Lecture, Oxford 1933.)

Introduction

Phonetics is the study of the sensible manifestation of language. It is, therefore, concerned with the acoustical properties of speech, with the motor behaviour of the vocal organs that produce the acoustical signal, and with the way the signal is processed in the human auditory system. In studying these topics sight must never be lost of the fact that we are dealing with manifestations of language, and not just with arbitrary acoustical signals produced by the human vocal tract and perceptible to our auditory system. Linguistic considerations and, in particular, facts concerning the limitations that different languages impose on the phonetic shape of words and utterances must, therefore, be taken into account on a par with articulatory and acoustic data. And while this has in fact been traditionally done, phoneticians have all too often done it surreptitiously, rather than explicitly and with good conscience. This is somewhat surprising since almost everything that phoneticians have investigated is crucially related to language.

The Notion of Segmentation

Fundamental to almost all research in phonetics is the concept of speech sound. It is the basic unit around which every textbook of phonetics has been organized from A. M. Bell's "Visible Speech" of 1867 to Peter Ladefoged's "Course in Phonetics" of 1975. In fact if the notion of speech sound were to be excluded from phonetic discourse—by decree of an all-powerful dictator—work in the field would come to a virtual standstill. Yet the notion of speech sound, i.e. the conception that utterances consist of discrete entities succeeding one another like the letters in a written text, is a purely linguistic notion that is not invariably given in the acoustical signal.

Consider, for example, the spectrograms of the Russian verb forms in Fig. 1

láyala "barked" (fem. sg.) láyu "I am barking".

While there are obvious changes in the resonance patterns with respect to time, these changes do not invariably occur abruptly so as to mark segment boundaries. For instance, the boundary between the [y] and the following [a] in *láyala* is far from immediately obvious. There can be little doubt, however, that the segmentation indicated in the phonetic transcription is indeed the correct segmentation as far as speakers of the language are concerned. To see this, consider a pair of verb forms quite similar to those in Fig. 1

s'iyála "shone" (fem. sg.) s'iyáyu "I am shining".

The two pairs of forms have the same morphological structure. In one case the stem is followed by the feminine singular past tense suffix [la]; in the other case, the stem is followed by the singular present tense suffix [u]. When we examine the stems, however, we notice at once that they differ. We have

$$\begin{array}{lcl} \text{láya} + \text{la} & \text{but} & \text{láy} + \text{u} \\ \text{s'iyá} + \text{la} & \text{but} & \text{s'yáy} + \text{u}. \end{array} \tag{1}$$

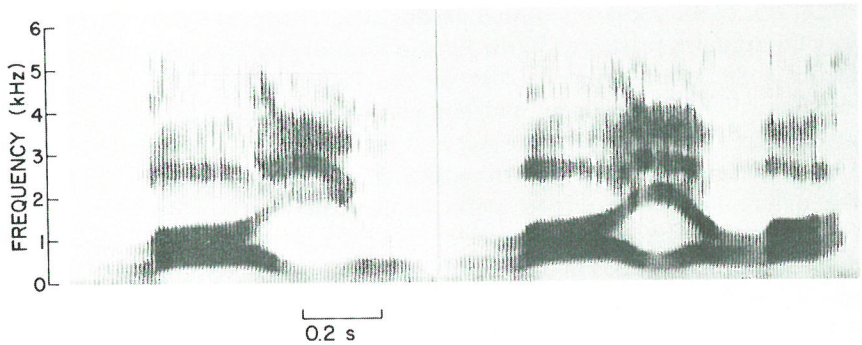

Fig. 1. Spectrograms of the Russian words *láyu* (left) and *láyala* (right).

The formula for the difference was discovered thirty years ago by Roman Jakobson (1948) who observed that Russian words were subject to the rule that

> stem final vowels are deleted before vowel-initial suffixes, while (certain) stem final consonants are deleted before consonant-initial suffixes. (2)

Thus,

láya + u → láy + u s'iyáy + u

láya + la s'iyáy + la → s'iyá + la. (3)

Proper application of Jakobson's rule implies that the speaker is able to segment an utterance such as [láya] at least into [lay] and [a], in spite of the fact that this segmentation is, as we have seen, not especially conspicuous in the acoustic signal. Since fluent speakers of Russian have full command of the Russian conjugation and, hence, by implication, obey Jakobson's rule in their utterances, they must also perceive (i.e. be capable of analysing) quasi-continuous acoustic signals such as those illustrated in Fig. 1 as sequences of discrete segments.* Since rules quite similar to (2) are a very essential part of the knowledge that underlies fluent command of every language that has ever been studied, segmentation of utterances into discrete sounds has always been taken as a language universal—a property possessed by all natural languages.

Classificatory Features

Phonological rules have another property that is of considerable interest here: they commonly involve not single sounds but groups of sounds, and these groups of sounds figuring in different rules in different languages not only exhibit striking overlaps, but are frequently also modified in the same fashion. We illustrate a few of these groupings.

(a) ɑ e i o u (stressed) i u ɑ (unstressed)
(b) ɑ ɔ ɛ o e i u (short) ā ɔ̄ ō ī ū (long)
(c) s z č ǰ š ž
(d) c ʒ ʒʼ s z č ǰ čʼ š ž̌. (4)

* We are able in the present instance to rule out the obvious suggestion that the segmentation is the result of the speaker's familiarity with the alphabetic writing system of Russian. On the one hand, in the instance under discussion Russian orthography obscures the fact that a single segment is being deleted, for the sequence [ya] is represented by a single letter of the Cyrillic alphabet. On the other hand, young children and adult illiterates speak fluent Russian, conjugate Russian verbs correctly, and show in their utterances that they obey the rules of Russian grammar (among them Jakobson's rule) in much the same fashion as speakers who have acquired literacy.

In the first set we have the vowels of Russian which are affected in rule (2). In (4b) we have the vowels of Dinka, which are involved in the plural formation rule to be discussed below (see 5). In (4c) we have the class of English consonants which is involved in the selection of the plural suffix [Iz] (see 6). Finally in (4d) we have the class of consonants of Navaho which are subject to the rule of consonant harmony (see Hoijer, 1945: 11–15). What is interesting about the different groups in (4) is that though they represent the same class of sounds—strident coronal consonants in (4c) and (4d), and vowels in (4a) and (4b)—their membership varies. Thus, the set of vowels in Russian is not identical with those of Dinka, nor is the set of strident coronal consonants identical in English and Navaho. However, linguists and phoneticians have traditionally and quite correctly treated such separate sets as different instantiations of the same phonetic classes, i.e. the class of vowels and the class of strident coronals, implying thereby that the respective groups are defined by sharing common features—vowelness in the former case, strident coronality in the latter.

It is a general characteristic of phonological rules that they view speech sounds as complexes of features. An instructive example to this effect is provided by Dinka, a language spoken in the Sudan. In Dinka, as already illustrated in (4b), there are two parallel sets of vowels, short and long. This fact is made use of in the formation of the plural of a certain class of Dinka nouns (see Gleason, 1955). In particular, in this class, nouns having a short stem vowel in the singular have a plural with a long stem vowel, and, vice versa, nouns with a long stem vowel in the singular have a short stem vowel in the plural.

pal-pāl "knife" čīn-čin "hand"

bit-bīt "spear" agōk-agok "monkey" (5)

In other words, everything proceeds here as though vowels were composed of at least two independent components: length and the rest. The plural formation rule affects only the length feature, leaving the rest intact. Moreover, the fact that long vowels alternate with short vowels and short with long implies that the classificatory feature of length (and we have argued elsewhere that this is also true of all features) is binary, so that "short" = "not long". The Dinka plural formation rule can then be stated as a "logical" negation of the length feature so that "long" becomes "not long" = "short", and "not long" becomes "not not long" = "long".

Facts such as these among others led Jakobson (1938) to propose that all speech sounds are complexes or "bundles of distinctive features". We want to suggest here that there is evidence for the psychological reality of this view of the sounds of speech; i.e. that the way the normal, phonetically and

linguistically unschooled speaker deals with speech sounds strongly suggests that he performs a distinctive feature analysis of them.*

As illustrated in (6a) English utilizes three distinct suffixes to form the plural of nouns. The choice of the three suffixes is governed by a rule which might be stated as in (6b).

(a) (i) [ɪz] busses, roses, leaches, ranges, bushes, garages
(ii) [s] caps, cats, cakes, fifths, coughs
(iii) [z] cabs, cads, cogs, cams, cans, calls, cars, lathes, plays, spies, blows, trees, . . . (6)

(b) The plural suffix is

(i) [ɪz] if the noun ends with [s z č j š ž], otherwise
(ii) [s] if the noun ends with [p t k f θ], otherwise
(iii) [z]. (6)

For each of the lists in (6b) we can substitute a feature bundle which will unambiguously designate the classes of sounds listed in (6b). The rule would then be reformulated as in (6c).

(c) The plural suffix is
(i) [ɪz] if the noun ends with a strident coronal sound, otherwise
(ii) [s] if the noun ends with a voiceless sound, otherwise
(iii) [z]. (6)

The question of interest is: which of the two formulations more accurately reflects the behaviour of real speakers? The appropriate experiment was suggested by Lise Menn (pers. comm.). It consists of presenting to speakers of English nouns that end with sounds lacking in the English consonant system and asking them to form their plurals. If they follow (6b), such "foreign" nouns should all receive the suffix [z], pursuant to option (6b, iii), which alone is applicable to such nouns. If, on the other hand, speakers utilize (6c) their response should take the feature composition of the sounds into consideration. A crucial test is provided by German names ending with [x] or [ç] such as *Bach*, *Reich*, *Milch*, *Kroch*. While we have not formally run the experiment, inquiries among our associates in the laboratory reveals that the plurals of these German names are commonly formed with [s] and never with [z], implying that phonetically unschooled speakers of English make use of formulation (6c) rather than (6b). As noted above, this implies further that phonetically untutored subjects analyse speech sounds into their classificatory features even when the sounds are demonstrably not part of their native

* The example below has been discussed previously in the literature (see Halle, 1977), but because of its great simplicity and instructiveness we utilize it here once again.

language. This analysis into features could not plausibly be said to have been learned, for there are surely few experiences in the life of a normal individual who is not a professional linguist or a phonetician that would lead her/him to develop a system of features for classifying speech sounds. One is, therefore, led to assume that the speech-analysing system is part of our genetic endowment, part of our uniquely human way of responding to the very special acoustic signals that make up an utterance in a natural language.

Evidence for Natural Classes of Speech Sounds from Examination of Articulatory Activity

If we examine the articulatory movements involved in speech production, for example by observing cineradiographic pictures of the speech-producing structures, it is not immediately evident that the apparently continuous sequence of movements can be segmented into smaller units corresponding to speech sounds. That is, we encounter difficulties in segmentation similar to those in the acoustic representation, as illustrated in Fig. 1. The process of extracting the underlying phonetic units becomes somewhat clearer, however, if we define a phonetic unit in articulatory terms as the achievement of a particular static target state or goal for the articulatory structures. That is, we view the process of speech production as consisting of the actualization of a sequence of target configurations or states of the articulatory structures. During most of the time in an utterance, the articulatory structures are moving between states, and only a small fraction of the time is spent at particular target states. Consequently, the description of the articulatory activity in terms of target states is not immediately evident from direct observation of articulatory movements. The articulatory movements and targets have certain other attributes that tend to obscure the representation of speech in terms of a sequence of discrete phonetic units.

(a) The sequence of target configurations or states is by no means uniformly spaced; sometimes the targets for two adjacent sounds are almost simultaneous, and at other times there is a substantial time interval between adjacent targets.
(b) A given target configuration may involve several articulatory structures and not all of these structures achieve their target states simultaneously; that is, there may be some asynchrony in the movements of the different structures such as larynx, velum, tongue body, tongue tip, etc.
(c) Situations arise in which the target configuration associated with a particular speech sound is not achieved before the articulatory structures proceed toward the next target; that is, there is undershoot in the articulatory movements towards some targets.

(d) The reaching of a particular articulatory target configuration does not necessarily register as an obvious acoustic event in the acoustic signal; the acoustic attributes that indicate the identity of a phonetic unit to a listener may occur at the time the target configuration is reached or may occur when the articulators are in transition between targets.

The target configuration or state that is used to produce a particular speech sound requires that each of several individual articulatory structures assumes a particular configuration. These different structures include the tongue body, the tongue blade, the soft palate, the laryngeal structures and the lips. If we examine the target configurations for the individual structures across a large number of speech sounds, we find that the same configuration or state for a given structure is used to produce many different sounds. Thus, for example, we observe that a number of sounds are produced with the blade of the tongue raised to make rather close contact with some part of the hard palate, although for different sounds within this group the other articulatory structures (such as tongue body, velum and larynx) may assume a variety of different configurations or states. This group of sounds can be considered to form a natural class by virtue of this common component of its articulatory description. Thus examination of the configurations or states for the individual articulatory structures when different sounds are produced provides us with a way of organizing speech sounds into classes.

We consider now several lists of phonetic segments. For all of the items on a given list, we suggest that some aspect of the articulation is achieving the same state, and that these items can therefore be considered to form a natural class.

(a) m n ŋ ã ũ
(b) k g ŋ i u
(c) p t k f θ s š á í ú
(d) p t k č b d g ǰ m n ŋ (7)
(e) p b f v m
(f) t d n θ ð s z š ž l r

The items listed in (7a) are all produced by creating a velopharyngeal opening, usually by placing the velum in a lowered position. From the point of view of the speaker, an indication that the velum is lowered comes from several possible sources

(a) the muscles used to lower the velum have been contracted,
(b) the lowered state of the velum is sensed through receptors that signal the position of the velum or its contact with other structures,
(c) there is airflow through the velopharyngeal opening and possibly acoustic energy in the nasal cavity that is sensed and registered in some way.

Any one of these three descriptions of the production of the sounds in (7a) (or of the orosensory consequences of the production) can be used to define the common articulatory state for this class of sounds.

The sounds in (7b) are all produced by placing the tongue body in a raised position within the oral cavity. More specifically, the common articulatory activity for these sounds can be described in one of two ways

(a) there is contraction of a common muscle or group of muscles to produce the raised tongue body

(b) there is a common pattern of activity in particular groups of sensory receptors in the tongue musculature or on the surfaces of the tongue as these surfaces make contact with other structures (Stevens, 1975; Dowla and Perkell, 1977).

For the group of sounds in (7c), it is hypothesized that the common articulatory attribute is a stiffening of the surfaces of the vocal folds. This stiffening leads to an inhibition of voicing when pressure builds up in the vocal tract behind the constriction of a consonant, and a raised fundamental frequency for vowels and other sonorants (Halle and Stevens, 1971). Again the articulatory state that characterizes each member of this class can be described either as contraction of a particular laryngeal muscle or group of muscles or as the state of the vocal fold surfaces, independently of the muscle activity used to produce that state.

The sounds listed in (7d) are all produced by forming a complete closure of the vocal tract at some point along its length. The articulatory description for this group of segments cannot be specified in terms of the contraction of particular muscles, since different muscles are clearly involved depending on where in the vocal tract the constriction is made. Rather, it is assumed that an instruction to form a complete closure is a basic component of articulatory control which, when coupled with a further instruction indicating which articulator is to be activated, effects the proper consonantal constriction. It is possible also that the sensory consequences of forming a complete closure are registered in some unique manner independently of the location of the closure in the vocal tract.

The segments listed in (7e) have the common articulatory attribute that they are produced with a constriction at the lips. Thus a particular set of muscles—those making a lip closure—is involved in the generation of all of these sounds. The lower lip comes in contact with either the upper lip or the upper incisors, and this gesture leads to a unique pattern of excitation of sensory units in the lower lip.

The phonetic segments given in (7f) are all actualized by raising the tongue blade to make contact with some part of the hard palate. The exact region of

contact or the force of contact may vary from one sound to another in the set, but the common gesture is that of raising the tongue blade, presumably through contraction of certain intrinsic tongue muscles. There is a unique sensory consequence of this raised position of the tongue blade: the edges of the superior portion of the tongue come in contact with fixed surfaces of the hard palate or teeth, presumably leading to a special response of tactile receptors on these surfaces of the tongue blade or of the palate.

The six lists of segments in (7) are examples of a longer inventory of lists of segments that could be generated. Furthermore, there is no attempt to make each list exhaustive; additional items could be appended to these lists. These examples serve to indicate, however, that natural classes of speech sounds can be constructed through examination of the articulatory target configurations or states. In giving these examples, we have shown a certain amount of ambivalence as to how the common articulatory attributes for the items on a list should be specified. On the one hand, it can be argued that in order to achieve the target configuration that characterizes the items on a list a particular pattern of muscle activity is required. On the other hand, the achievement of the target configuration or state results in a unique pattern of response of sensory units located within an articulatory structure. Arguments for utilizing common patterns of orosensory responses to characterize natural classes of speech sounds have been presented elsewhere (Stevens and Perkell, 1977). Furthermore, the articulatory attribute that characterizes a particular class of sounds does not always refer to a specific articulatory structure, but instead may refer to a type of gesture, such as making a complete closure, that could be actualized using any one of several structures. Until we know more about how motor systems operate, and, in particular, how the speech-production systems operate, the question of how best to characterize natural classes of speech sounds in terms of articulatory attributes must remain open.

Acoustic and Psychoacoustic Evidence for Natural Classes

Acoustic analysis of speech shows that there are groups of speech sounds that seem to have common acoustic properties. If it is assumed that the auditory system responds in some unique way to sounds with a common acoustic property, then this unique response provides the listener with a means of organizing speech sounds into natural classes based on their acoustic properties.

As examples, consider the lists of speech sounds given (8) below. Basing our observations on the work of Gunnar Fant (1960), of Jakobson *et al.* (1963), and of others, we shall show that for the items in any one of these lists there is a

common acoustic property, and that this acoustic property is different from one list to another.

(8)

(a) m n ŋ
(b) t d n
(c) k g ŋ
(d) i ɪ u ʊ
(e) ã ũ ĩ
(f) p t k č b d g ǰ m n ŋ
(g) á ú í
(h) p t k f θ s š č

For the items in (8a), there is a rather steady nasal murmur persisting for several tens of milliseconds. The unique acoustic attribute of this nasal murmur is a strong spectral peak at low frequencies and a relatively uniform distribution of weaker spectral peaks at higher frequencies, with these peaks tending to be rather broad (Fujimura, 1962).

In the case of the consonants in (8b), the spectrum sampled at the consonantal release (in a consonant-vowel syllable) shows a diffuse spread of energy across the frequency range, but with greater spectral energy at high frequencies. The sounds listed in (8c), on the other hand, have an onset spectrum with a single prominent peak in the midfrequency range (Zue, 1976; Blumstein and Stevens, in press).

The vowels listed in (8d) all have a relatively low first formant. The nasalized vowels in (8e) have a spectrum in which the lowest peak, corresponding to the first formant region for a nonnasal vowel, is split or broadened to cover a wider frequency range than that for a nonnasal vowel.

The items in list (8f) all show an abrupt onset of spectral energy over much of the frequency range when the consonant is released into the following vowel. The rise in spectral energy in any one frequency band occurs in a time interval of just a few milliseconds. A sound with an abrupt onset has been shown to produce a distinctive response in a listener (Cutting and Rosner, 1974). The vowels listed in (8g) all have a fundamental frequency (F_0) that is high in comparison with the average F_0 for the particular speaker and the particular position of the vowel within an utterance. The common acoustic characteristic of the sounds listed in (8h) is the absence of low-frequency periodicity in the sound in the vicinity of the consonantal closure interval.

On the Relationship among Features in the Different Domains

The evidence presented has shown that in each of three different domains speech sounds should be viewed as complexes of attributes rather than as atomic, further unanalysable entities. The question that naturally arises at this

point concerns the relations between the attributes in each of the three domains.

If we examine the lists of items in (7) and (8), we observe that the groupings of segments for the lists have a great deal in common. This is not too surprising in view of the fact that there are well-established relations between the configuration of the vocal tract and the sound resulting from this configuration. Much of our understanding of the relationship between articulation and acoustics derives from the pioneering researches of Gunnar Fant. (See, in particular, the classic study in Fant, 1960.) There is not always, however, a one-to-one relation between the classes of segments based on articulatory attributes and the classes based on acoustic attributes. List (7d) and list (8f) provide an example of what appears to be a one-to-one relation between the two groupings. Consonants produced with complete closure of the vocal tract always show an abrupt onset of acoustic energy at the release of the closure. On the other hand, the items in lists (7a) seem to be divided into two lists in (8): lists (8a) and (8e). At least based on a superficial analysis of the acoustics, nasalization seems to be manifested in the sound in different ways for consonants and for vowels. And even more complicated instances can readily be found.

However our principal concern here is not only with the relationship between the articulatory and acoustical attributes of language but rather how these attributes relate to the classificatory features. The classificatory features constitute the most direct link between the rules of language and the physical and physiological facts of speech. It is obvious that the relationship between classificatory features and the acoustical and articulatory attributes cannot, in general, be inferred from the relationship that exists between articulation and acoustics. Phonetic theory must, therefore, include a special chapter listing the articulatory and acoustical correlates of each classificatory feature.

Since phonetic theory must pay detailed attention to the classificatory features, the arguments and reasons for the different features are of vital concern to the discipline. The major classificatory evidence in favour of a given feature derives from groupings of sounds in different rules and from the modifications that such sounds are subject to. It is to be expected that facts from this domain will also shed light on articulatory and acoustical properties of language.

An example of this interaction between classificatory and articulatory and acoustical considerations is provided by recent discussion concerning the feature coronal. In Chomsky and Halle (1968) (SPE) the class of coronals was said to include dental, alveolar, and palato-alveolar consonants and liquids, whereas "the uvular [R] and the consonants articulated with the lips and with the body of the tongue are noncoronal" (SPE: 304). Specifically included among the latter were palatal consonants such as the German [ç] (*ich*) sound

and the glide [y]. More recently a number of papers have appeared challenging this classification. Thus, Hyman (1973), Vago (1976) and Odden (1978) have adduced a number of phonological rules from various languages where palatal consonants and the front glides [y, ẅ] must be grouped with the coronals. Odden cites, among other facts, the well-known Sanskrit rule of n-retroflexion which applies whenever [n] is preceded by [ṣ, ṛ, r] anywhere in the word, provided only that no dental, alveolar, or palatal consonant intervenes, while Vago draws attention to a rule of Baule, a language of the Ivory Coast, where "[w] is pronounced as the front glide [ẅ] between an alveolar or palatal consonant and a front vowel". The variety of the rules and the wide geographic distribution of the languages shows that we are not dealing here with a few isolated instances but with a perfectly common phenomenon i.e. palatal consonants and front glides are grouped together in a single phonological class with dental, alveolars and palato-alveolars.

An immediate consequence of the recognition of the modification of the classificatory feature coronal is the need to state the articulatory and acoustic correlates of the newly revised feature. In SPE, coronal sounds were said to be "produced with the blade of the tongue raised from its neutral position," thus specifically excluding palatal consonants, which are produced by raising not the blade of the tongue, but rather its central portion, i.e. the part connecting the blade with the tongue body. The articulatory correlate of the revised feature of coronality is, therefore, the raising of the frontal (i.e. tip, blade and/or central) part of the tongue so as to make contact with the palate.

If palatal consonants are to be classified as coronals, it becomes necessary to specify how they are to be distinguished from the other coronals. In effect, we need to distinguish between the postalveolar fricative consonants [ç ṣ š], and there may be parallel distinctions for the nasals and for the stops. With regard to the nasal and stop counterpart of the [š], the literature is somewhat unclear, however, and we shall limit our comments to the fricatives. When we examine the articulatory mechanisms involved in the postalveolar fricatives, we note that an essential aspect of these sounds is whether or not the front of the tongue is raised up in such a way as to affect the direction and character of the air stream. Some phonetics texts (e.g. Smalley, 1977: 167) characterize this distinction as one between a grooved and flat shape of the tongue. This is the major difference between the articulation of the German [š] and [ç] in *Tisch* "table" versus *dich* "you". The raised apex [š] directs the air stream to impinge on the upper teeth or on some other obstruction anterior to the constriction, whereas the lowered apex of [ç] prevents the teeth from being struck by a narrow stream of air. What distinguishes these two sounds from the third postalveolar coronal, the retroflex [ṣ], is, as proposed in SPE, "the presence of a constriction that extends for a considerable distance along the direction of the air flow". In the terminology of SPE, [ç] and [š] are [+distributed], while

the retroflex [ṣ] is [−distributed]. As noted above, what distinguishes the two kinds of distributed fricatives is the property that we designate tentatively with the term grooved: [š ṣ] are grooved, whereas [ç] is nongrooved or flat. The fourth logically possible category—[−distributed, −grooved]—does not seem to exist.

We observe the same three-way contrast in the dental region. Thus, the interdental [θ] is produced with a flat tongue, and the air stream is not directed against the upper teeth. Grooving is, however, involved in the ordinary English [s]. The third class of anterior consonants is represented by grooved dentals produced with a constriction that extends for a considerable length along the teeth ridge. One finds such consonants in some languages (e.g., Spanish) as the basic variant of [s], whereas in English it appears to be a free variant of the apical [s] (see Jones, 1960: 185–186).

In summary, the revised feature coronal leads to the introduction of a new feature grooved, yielding the following feature characterizations of the major classes of consonants under discussion here (the feature anterior is used here in the same sense as in SPE):

	θ	s	s_1	ç	ṣ	š
anterior	+	+	+	−	−	−
distributed	+	−	+	+	−	+
grooved	−	+	+	−	+	+

(9)

It may be observed, moreover, that the feature grooved appears to make the distinction between strident and nonstrident rather marginal, if not altogether superfluous. For instance, we now can characterize the class [s z č j š ž] which plays a role in the English plural rule (cf. 4) as [+coronal, +grooved]. Further study will, however, be required before this further modification of the feature system can be formally proposed.

Acoustically, the revised feature coronal is characterized by a gross spectrum shape that shows a greater concentration of energy at high frequencies than at low frequencies. The greater energy at higher frequencies is a consequence of the spectrum of the noise which is produced in the relatively short acoustic cavities at the constriction and anterior to it in the case of obstruent coronals and the raised frequencies and the proximity of the second and higher formants in the case of the glide (Hughes and Halle, 1956; Fant, 1960). All of the consonants in (9) appear to have this general property.

The acoustic correlate of the feature grooved has yet to be specified in detail. One factor is that the source of noise is in the vicinity of the teeth for grooved fricatives, but is located at the tongue constriction for the nongrooved fricatives. This different source location leads to a different overall spectrum balance at high frequencies, as has been shown theoretically by Fant (1960: 177–178). Furthermore, the frication noise in consonants with this feature

seems to have greater overall intensity than in the nongrooved, and there appears to be a different character to the noise, probably due to more dominant peaks in the noise spectrum for the apicals. The greater intensity is a consequence of the more efficient generation of noise when the turbulence that produces this noise results from a rapid air stream impinging on an obstruction such as the teeth. The presence of narrower, more dominant spectral peaks may arise from the fact that the posterior termination of the acoustic cavity in front of the constriction is rather abrupt in the case of the grooved configuration shaped by the tongue tip, whereas this termination narrows more gradually in a posterior direction for the nongrooved configuration, leading to greater acoustic losses (and hence a greater bandwidth for the front-cavity resonance). This question needs further study, however.

We have been led to redefine a classificatory feature and to introduce a new feature on the basis of phonological observations from language. This need for slight modification of feature definitions arising from linguistic evidence then led to a reexamination of the acoustic and articulatory evidence. On the basis of this reexamination it has been possible to suggest hypotheses regarding the acoustic and articulatory bases of the newly defined features, and these hypotheses in turn suggest what experimental studies are needed in order to substantiate or to disprove them. The exercise, then, is an example of the interplay between acoustics, psychoacoustics, articulatory studies and phonology that will contribute to a more firm basis for a system of distinctive features.

Acknowledgements

This work was supported in part by the National Institutes of Health, grant NS-04322 and the National Institutes of Mental Health, grant MH-13390.

References

Blumstein, S. E. and Stevens, K. N. (in press). Acoustic invariance in speech production: Evidence from measurements of the spectral characteristics of stop consonants. *J. Acoust. Soc. Am.*

Chomsky, N. and Halle, M. (1968). "The Sound Pattern of English." Harper and Row, New York.

Cutting, J. and Rosner, B. (1974). Categories and boundaries in speech and music. *Percept. Psychophys.* **16,** 564–570.

Dowla, F. and Perkell, J. S. (1977). Quantal nature of speech production for vowels: Implications of cineradiographic and palatographic observations. *J. Acoust. Soc. Am.* **62,** S15.

Fant, C. G. M. (1960). "Acoustic Theory of Speech Production." Mouton, The Hague.
Fujimura, O. (1962). Analysis of nasal consonants. *J. Acoust. Soc. Am.* **34,** 1865–1875.
Gleason, A. (1955). "A Workbook in Descriptive Linguistics." Rinehart and Winston, New York.
Halle, M. (1977). Language and communication. *In* "The Telephone's First Century and Beyond." (Clarke, A. C. *et al.*, eds). New York: Crowell, New York, pp. 39–59.
Halle, M. and Stevens, K. N. (1971). A note on laryngeal features. Research Laboratory of Electronics Quarterly Progress Report, No. 101, MIT, pp. 198–213.
Hoijer, H. (1945). "Navaho Phonology." University of New Mexico Press, Albuquerque.
Hughes, G. W. and Halle, M. (1956). Spectral properties of fricative consonants. *J. Acoust. Soc. Am.* **28,** 303–310.
Hyman, L. (1973). The feature [grave] in phonological theory. *J. Phon.* **1,** 329–337.
Jakobson, R. (1938). Observations sur le classement phonologique des consonnes. Proc. 3 Int. Congr. Phon. Sci., Ghent, pp. 34–41.
Jakobson, R. (1948). The Russian conjugation. *Word* **4,** 155–167.
Jakobson, R., Fant, C. G. M. and Halle, M. (1963). "Preliminaries to Speech Analysis." MIT Press, Cambridge, USA.
Jones, D. (1960). "An Outline of English Phonetics." W. Heffer & Sons, Ltd, Cambridge.
Odden, D. (1978). Further evidence for the feature [grave]. *Ling. Inquiry* **7,** 671–674.
Smalley, W. (1977). "A Manual of Articulatory Phonetics." William Carey Library, Pasadena.
Stevens, K. N. (1975). Quantal configurations for vowels. *J. Acoust. Soc. Am.* **57,** S70.
Stevens, K. N. and Perkell, J. S. (1977). Speech physiology and phonetic features. *In* "Dynamic Aspects of Speech Production" (Sawashima, M. and Cooper, F. S. eds). University of Tokyo Press. Tokyo, pp. 323–341.
Vago, R. M. (1976). More evidence for the feature [grave]. *Ling. Inquiry* **7,** 671–674.
Zue, V. W. (1976). Acoustic characteristics of stop consonants: A controlled study. Ph.D. thesis (unpublished), MIT.