UCLA
Working Papers in Phonetics

Title
WPP, No. 6: Linguistic Phonetics

Permalink
https://escholarship.org/uc/item/40j0x60d

Author
Ladefoged, Peter

Publication Date
1967-06-01

eScholarship.org
Powered by the California Digital Library
University of California

LINGUISTIC PHONETICS

PETER LADEFOGED

WORKING PAPERS IN PHONETICS 6

UCLA
JUNE 1967

L I N G U I S T I C   P H O N E T I C S

Preliminary version for comment and criticism

P e t e r   L a d e f o g e d

Preface	1
Chapter 1. The phonetic bases of phonological contrasts	3
Chapter 2. Preliminaries to a theory of phonetics	55
Chapter 3. Universal phonetic parameters	74
Bibliography	90
Index of languages cited	94
General index	96

Working Papers in Phonetics, 6. June 1967

Phonetics Laboratory, University of California, Los Angeles

Typographical errors corrected, and minor

changes made September 1967

PREFACE

This book presents a possible set of phonetic categories for describing the languages of the world, and an account of their internal organization. It is thus a statement of part of the theory of general linguistics. The reader is assumed to have read at least an introductory text in general linguistics, such as Gleason (1961); and he may find it convenient to know more of the phonetic terminology through having read one of the introductory phonetic texts, such as Abercrombie (1967). Fischer-Jørgensen (1960), Heffner (1950), Malmberg (1963), or the early critical survey of the field by Pike (1943). He is also assumed to be familiar with the symbols of the International Phonetic Association (1949).

The descriptions of sounds given here are all based on my own observations with informants, unless a source is explicitly cited. But I have usually relied on other people for the phonological description of each language, contenting myself with trying to describe the phonetic nature of the contrasts which they have determined. Thus an unqualified statement to the effect that a certain language (e.g. Marwari or Margi) distinguishes between certain pairs of sounds (respectively h-ɦ and l-ɮ) implies that I have heard native speakers of these languages making the distinctions in the way described; and that these sounds contrast, at least superficially. (But it does not, of course, follow that these contrasts should necessarily be represented in the underlying phonological forms.)

Because of this predilection for basing my remarks on personal observation, what I have written is often inadequate in that my data are from a very limited number of speakers and a very small proportion of the languages of the world. I have investigated the phonetic structure of a number of African, Indian and European languages, and have been able to make scattered observations of some of the prominent languages (and a few of the less well known ones) outside these areas. But there are many language families which I know only through the literature; and, as most practising phoneticians would agree from experience, published phonetic descriptions are often impossible to interpret accurately. I hope the present formulation of a theory of phonetics will help improve this situation. But the statements made here should all be regarded as tentative; they are made in the belief that "a rule requiring amendation is more useful than the absence of any rule" (Jakobson 1962). I would be grateful for comments on the work as a whole,

and corrections of the statements about particular languages and the existence or non-existence of particular combinations of categories. I have already profited greatly from the comments of my colleagues, Victoria Fromkin and William Bright, and many of the students (or former students) at UCLA, including especially Frank Heny, Kalon Kelley, Chin Kim, John Ohala, Tim Smith, Marcel Tatham, Ralph Vanderslice, and Harry Whitaker.

This publication is one of an irregular series of *Working Papers in Phonetics* put out by the UCLA Phonetics Laboratory. Much of the work reported in this issue is supported through an NSF grant for research on Linguistic phonetics. A grant from UCLA funds has made possible the publication of the MS in its present form, so that it might be made available to research workers in the field for their comments. The MS undoubtedly contains many errors, and is not yet ready for regular publication.

1

THE PHONETIC BASES OF PHONOLOGICAL CONTRAST

Whenever we describe a language, at some point we have to talk about the sounds; and when we do this, we normally use phonetic categories, such as voiced, alveolar, and fricative (or perhaps acute, and strident). One task for phoneticians is to provide categories which are adequate for linguistic purposes and which can be correlated with extra-linguistic events describable in physiological, physical terms. Obviously part of this task involves being able to specify the linguistic oppositions (such as the phonemes) which occur within languages. In the first part of this monograph we will describe the speech mechanism using a traditional approach based on that of Pike (1943) and Abercrombie (1967), and at the same time discuss the way in which the mechanism is used to produce many of the systematic phonetic contrasts in the languages of the world. We will then be in a better position to assess some of the data which a theory of linguistic phonetics must encompass.

We may consider speech to be the product of four separate processes: the airstream process, the phonation process, the oro-nasal process, and the articulatory process. As a first simplification we may associate these four processes with the actions of the lungs, the vocal cords, the velum, and the tongue and lips, as shown schematically in Figure 1; but, as we shall see later, the act of moving a body of air (which is termed the airstream process) may be more complicated. Variations in the airstream process which produce linguistic contrasts will be discussed after we have considered the actions of the vocal cords (which are summed up under the label of the phonation process) and the states of the velum (which form the oro-nasal process). This chapter will then be completed with an account of those linguistic contrasts dependent on the actions of the tongue and lips which form the articulatory process.

The Phonation Process

In a somewhat speculative article on the phonation process Catford (1964) describes more than ten states of the vocal cords which are linguistically significant. The evidence for some of these states is a little vague, as Catford himself is well aware, explicitly stating that his article is only a preliminary attempt at setting up a scheme of categories for types of phonation. But most phoneticians would agree that we probably need at least six

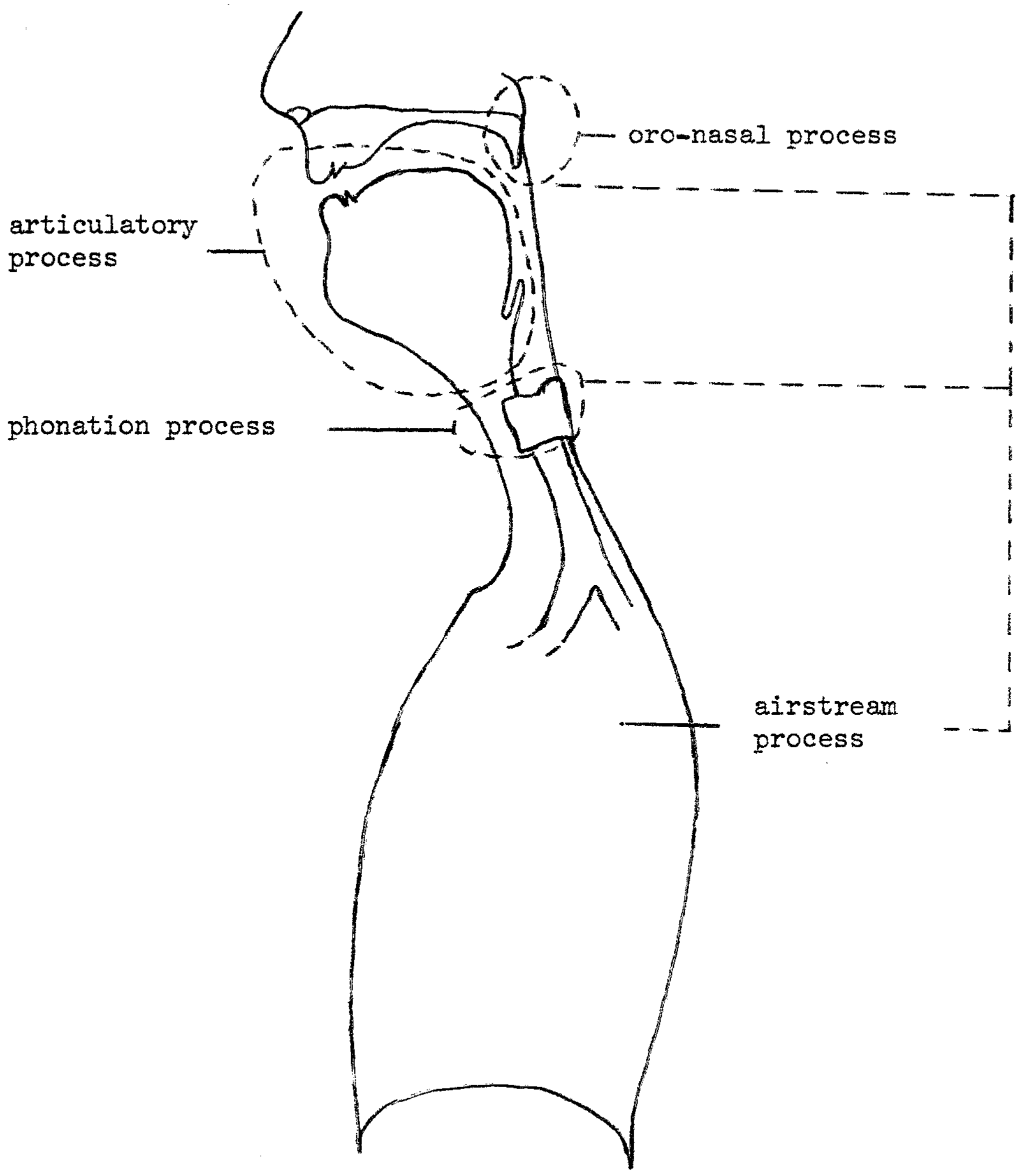


Figure 1: Schematic diagram of the vocal organs, showing the four processes required in the specification of speech.

or seven states of the vocal cords to account for the linguistic oppositions which occur within languages. These states are summarized in Table 1.

The positions in voiced and voiceless sounds are too well known to need much further comment here. In the formation of a voiced sound the vocal cords are adjusted so that they are almost touching along their entire length. The result of air flowing through this constriction is a suction effect which draws the vocal cords together. But as soon as they are together there is no flow, and consequently no action pulling them together; so they come apart and release the pressure which has been built up beneath them. But when they are apart they are again subject to the suction caused by the outgoing air. So the cycle repeats itself, producing the regular vibrations known as voice.

The rate of vibration during a voiced sound depends on two factors: the tension of the vocal cords; and the pressure drop across them. The pitch will go up whenever either of these two factors is increased, as explained elsewhere (Ladefoged 1967). The forces exerting tension on the vocal cords can also be divided into two groups, in the first being the muscles of the larynx, and in the second the articulatory movements of the tongue which tend to raise the larynx and hence stretch the vocal cords. Changes in pitch which are linguistically significant are due to either changes in the pressure of the air below the vocal cords, or changes in the tension of the vocal cords caused by the action of the muscles of the larynx. In all languages so far observed, all other changes in pitch have been shown to be predictable. Thus in a tone language the actual frequency of vibration during the occurrence of a particular tone may be slightly higher in a syllable containing the vowel i , in which the larynx tends to be pulled upwards by the high tongue position, than in a similar syllable containing a more open vowel such as ɑ . In an intonation language such as English the frequency of the peak of an intonation contour depends in a similar way on the vowel quality (Lehiste and Peterson 1961). Neither of these variations are linguistically significant.

We do not yet know which factor is most important in making significant pitch changes. The intonation contour in some statements in English has a high correlation with the tension of the vocal cords (Ladefoged, 1963 and 1967). But it appears (from hitherto unpublished material) that the high pitch which occurs at the end of tag questions such as 'It's true, isn't it?' may also be associated with an increase in the pressure of the air below the vocal cords in many speakers. Other recent unpublished studies indicate that in a tone language such as Yoruba there is often an increase in subglottal pressure during high tones.

During voiceless sounds the vocal cords are apart. They are not, however, pulled as far apart as possible. In normal expiration they are slightly closer together than they are in inspiration, and as far

Table 1: Some phonation types

Phonetic term	Brief description	Example symbols
Voice	vibration of the vocal cords	m,z,b,a
Voiceless	vocal cords apart	m̥,s,p,h
Aspiration	a brief period of voicelessness during and immediately after the release of an articulatory stricture	s^h,p^h
Murmur	'breathy voice' -- arytenoids aparts, ligamental vocal cords vibrating	m̤,z̤,b̤,
Laryngealization	'creaky voice' -- arytenoids tightly together, but a small length of the ligamental vocal cords vibrating	m̰,z̰,b̰
Glottal stop	vocal cords held together	ʔ
Whisper	vocal cords together or narrowed except between the arytenoids	(no symbol)

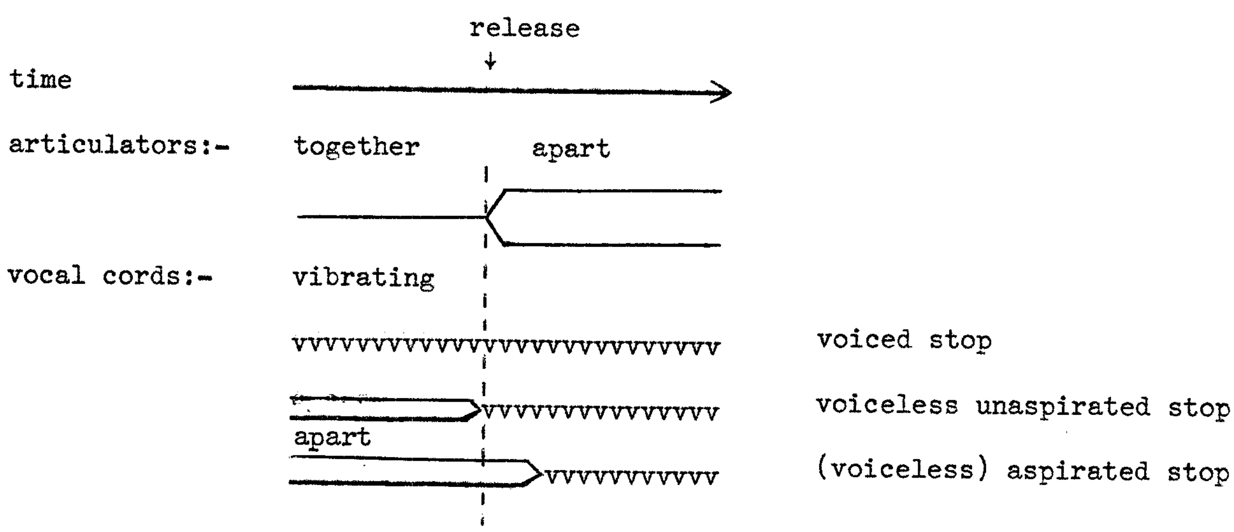


Figure 2: The relation between the timing of the articulatory movements and the state of the glottis in some stop consonants. Note that the terms voiced - voiceless refer only to the state of the glottis during the articulatory stricture, and the terms aspirated - unaspirated only to the state of the glottis during and immediately after the release of the stricture.

as is known the position during voiceless sounds is the same as that during expiration. This difference between inspiration and expiration may be why h sounds are sometimes said to have a slightly narrowed position of the vocal cords; but the position for an h which is not between voiced sounds is probably the same as in any other voiceless sound. In fact, h is often simply a phonologically convenient way of designating a sound which is the voiceless counterpart of the adjacent, usually the following, sound.

Many linguistic contrasts can be characterized simply by means of the oppositions voiced - voiceless, and aspirated - unaspirated. We shall use all these as technical terms with precise definitions. Voiced and voiceless refer to specific states of the vocal cords occurring during the articulation of a sound. If a sound is said to be partially voiced (or voiceless), the state of the vocal cords is that for voice (or voicelessness) during only part of the articulation. English initial b may be said to be usually partly voiced, whereas French initial b is nearly always fully voiced. Aspirated and unaspirated refer to the state of the vocal cords during and immediately after the release of an articulatory stricture. In any aspirated sound the vocal cords are in the voiceless position during this period; in an unaspirated sound they are in a voiced position. In a voiceless unaspirated sound they start vibrating at about the same time as the stricture is released; and in a voiced sound they vibrate during both the closure and its release. The three possibilities are schematized in Figure 2. It should be noted that the fourth possibility, a voiced aspirated sound would, from the standpoint of these definitions, be a sound in which the vocal cords were vibrating during the articulation and then came apart into the voiceless position during the release of the stricture. Such a sound has not yet been observed in any language; we will consider later the structure of sounds such as Hindi bɦ, dɦ, gɦ which are often called voiced aspirates.

The definition of aspiration found in American phonetic literature sometimes refers to the release of an extra puff of air. This usage is rejected here because it is not specific enough. There are at least two possibilities as to how this extra puff of air might be produced: it could be the result of an extra push from the respiratory muscles; or it could be due to a valve-like action of the glottis allowing more air to be released. If we want our phonetic descriptions to embody a precise account of how a speaker makes a sound we should avoid a usage which permits the specification of two different actions which in fact produce different sounds.

Contrasts between voiced and voiceless versions of nearly all types of sounds occur. Examples of languages having simple contrasts between voiced and voiceless stops and fricatives need not be given here. Voiced and voiceless nasals and laterals occur in a number of languages;

Table 2: Voiced and voiceless nasals and laterals in Burmese

mà	healthy	nà	pain	ŋa	fish
m̥à	order	n̥à	nostril	ŋ̊á	rent
		lá	moon		
		l̥á	beautiful		

Table 3: Aspirated, voiceless unaspirated and voiced stops in Thai

pʰàa	to split	tʰam	to do	kʰàt	to interrupt
pàa	forest	tam	to pound	kàt	to bite
bàa	shoulder	dam	black		

Table 4: Voiced, aspirated, and unaspirated voiceless fricatives in Burmese

zãn	levitation	zauŋ	edge
sãn	example	sauŋ	harp
sʰãn	rice	sʰauŋ	winter

Table 5: Murmured stops and vowels in Gujerati, showing both the phonemicization suggested by Pandit (1957) and a transcription indicating the phonetic segments

/bar/	[baɾ]	twelve	/pɔr/	[pɔɾ]	last year
/bahr/	[ba̤ɾ]	outside	/pɔhr/	[pɔ̤ɾ]	early morning
/bhar/	[b̤aɾ]	burden	/phɔdz/	[pʰɔdz]	army
/aṛ/	[aṛ]	obstruction			
/ahṛ/	[a̤ṛ]	bones			

examples of contrasts in Burmese are given in Table 2. An acoustic analysis of the consonants at the beginning of the first group of words as said by two informants showed that voicing began shortly before the closure was released in each case. I do not know if this is typical of all speakers and all languages having contrasts of this kind, but if it is, perhaps these sounds should be called partially voiced as opposed to fully voiced. Voiceless semivowels such as those in Scottish English *which* (as opposed to *witch*) also have the voicing starting before the steady state of the vowel has begun; but it is largely an arbitrary question as to whether this point should be taken as marking the end of the semivowel. Voiceless vowels occur as allophones in many languages. According to Lounsbury (personal communication) none of the cases cited in the literature on Amerindian languages actually constitute examples of phonemic contrasts between voiced and voiceless vowels.

Contrasts between aspirated and unaspirated sounds are common among stop consonants. Many languages have three way oppositions as exemplified by the Thai words in Table 3. Aspirated and unaspirated fricatives also occur, but three way contrasts are less common; Burmese examples are given in Table 4. As far as I know no language uses a contrast between aspirated and unaspirated semivowels, although it is perfectly easy to distinguish between sequences such as ʍ^{h}a and ʍa . Sounds in which a period of voicelessness occurs before and during the formation of a stricture are said to be preaspirated; unit phonemes of this kind occur in Gaelic and Icelandic.

Many sounds cannot be characterized in terms of the two states of the vocal cords, voiced and voiceless. In Gujerati, and in several other Indian languages, there is an opposition between two sets of vowels, in both of which the vocal cords are vibrating. Firth (1957) described the one as having tight phonation and the other breathy phonation. I prefer to follow Pandit (1957) in referring to one as voice and the other as murmur. In the one set I can find no difference from the kind of vibrations of the vocal cords described above as voice. The other set is 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. There is a high rate of flow of air out of the lungs during these sounds, so the term breathy voice is also quite appropriate. English ɦ between vowels (as in *ahead*) has a similar quality.

In Gujerati this third state of the vocal cords, which is quite different from that for voiced or voiceless sounds, may not be accompanied by an extra push of the respiratory muscles; the different mode of vibration of the vocal cords is probably due to a different adjustment of the larynx. The examples in Table 5 show that murmur can occur during both consonants and vowels (or, as Firth might have put it, it can be a prosody of either syllable initials or syllable finals). There is no agreed IPA diacritic for marking this kind of phonation. I have used a subscript umlaut.

The phonemicization suggested by Pandit (1957) is also shown. Gujerati does not have a contrast between voiced and murmured vowels after murmured consonants

Most Indo Aryan languages have a series of stops with a murmured release, in addition to a three way contrast between voiced, voiceless unaspirated, and (voiceless) aspirated stops as described above. In all the languages I have examined (Hindi, Sindhi, Marathi, Bengali, Assamese, Gujerati, Bihari, Marwari, and others) the murmured stops are clearly distinguished by having a different mode of vibration of the vocal cords. There are minor variations (Saurashtri, as spoken at Madurai, sounds as if the release of the stops is accompanied by a brief period of voice before the murmur), but in every case one cannot construct a model which will generate these sounds without allowing for three distinct states of the vocal cords. It was for this reason that the murmured sounds could not be fitted into the scheme of Figure 2, which symbolizes only two different states of the vocal cords. There is, it is true, an extra puff of air accompanying both the voiceless aspirated and the murmured stops; but this puff of air is produced in a different way in each case so that the release of the one sounds quite different from the other. Phonemically it may be very convenient to symbolize these sounds as b , bh , p , ph etc; but when one uses a term such as voiced aspirated, one is using neither the term voiced nor the term aspirated in the same way as in the descriptions of the other stops. Murmured stops could be represented on a diagram like Figure 2 only by using a different kind of line to represent a third possible state of the vocal cords.

Murmured consonants occur in a number of languages outside India. They are common in Southern Bantu languages, such as Shona, Tsonga and members of the Nguni group. In all these languages during the murmured sounds the vocal cords seem to be held slightly closer together than in the Indian languages, so that there is more voice and less breath escaping; nevertheless they contrast clearly with the mode of vibration of the vocal cords which occurs in regular voiced sounds. In Shona there are voiced and murmured nasals. In the Nguni languages Ndebele and Zulu the situation is somewhat similar. There are contrasting voiced and murmured nasals like those in Shona. These sounds, together with a type of *h* which is realized as a murmured onset of a vowel, and the stops written *b, d, g* (and perhaps some other consonants) form a phonological class recognizable by the fact that they may cause a noticeable lowering of the tone on the subsequent vowel. Phonemically they may be considered to be mh , nh , hh , bh , dh , gh contrasting at least with m , n , h , g . But although the difference between mh , nh and m , n is that between murmur and voice, and the difference between hh and h is that between murmur and voicelessness, the difference between gh and g seems to be simply that gh is a depressor of the tone on the following vowel; both seem to have ordinary voicing. There are thus two voiced velar stops which are phonemically distinct only because of their influence on

Table 6: Voiced and murmured nasals, murmured and voiceless approximants (semivowels), and depressor and non-depressor stops in Ndebele

/úmúntu/
[úmúntu] person

/úgúna/
[úgúna] to run

/úmhámha/
[úm̤âm̤a] my mother

/úmnháwámi/
[úmn̤âwámi] my young brother

/úgúhhúla/
[úgúɦṳ̂la] to be a prostitute

/úgúhámba/
[úgúhámba] to travel

/égúvugéni/
[égúvugéni] on getting up

/égúgughéni/
[égúgugêni] on growing old

/gh/ = [depressor-of-following-high-tone-g]

Table 7: Contrasts involving laryngealized stops and semivowels in Margi

bábál	open place	b̰àb̰àl	hard
bʷál	ball	b̰ʷàb̰ʷà	cooked
b͜dàgò	valley	b̰d̰èbdò	chewed
dàlmà	big axe	d̰àd̰àho	bitter
káwà	sorry	w̰áw̰í	adornment
jà	give birth	j̰à	thigh

Table 8: Contrasts involving laryngealized vowels in Lango

lee	animal	lḛḛ	axe
man	this	ma̰n̰	testicles
kor	chest	ko̰r̰	hen's nest
tur	break	tṵr̰	high ground

the tonal pattern. Ndebele examples (based on Fortune, 1966) are given in Table 6

Another mode of vibration of the vocal cords occurs in laryngealized sounds. In this type of phonation the arytenoid cartilages are pressed inwards so that the posterior portions of the vocal cords are held together and only the anterior (ligamental) portions are able to vibrate. The result is often a harsh sound with a comparatively low pitch. It is also known as vocal fry (Moore and von Leden 1958) and creaky voice. Catford (1964) distinguishes between creak and creaky voice, but I am not sure if this distinction is needed for a theory of linguistic phonetics.

The opposition between voicing and laryngealization occurs during both semivowels and stops in Chadic languages such as Hausa, Bura and Margi. Something similar occurs in nasals in Amerindian languages such as Navaho. Laryngealized phonation is indicated by a subscript tilde in the Margi examples cited in Table 7. In these words, as in similar forms in related Chadic languages, laryngealized voicing is often audible in the adjacent vowels; but the laryngealization is regarded as a feature of the consonant not only on the phonetic grounds that it is clearly more evident during the consonant, but also on the distributional grounds that it occurs on all vowels, but only when adjacent to these certain consonants. In Nilotic languages such as Ateso and Lango there is another form of laryngealization which is used to distinguish a set of five vowels from a similar set with more normal voicing. Examples are given in Table 8. These languages have significant tones which can occur with either kind of phonation. The auditory difference between the two sets of vowels is that between a harsher, more reverberant, sound, and one with a softer voicing, nearer to (but not the same as) murmur. There is no report of a languages making use of the distinction between laryngealization and murmur, which may lead us to a different way of looking at the phonation types murmur, voice, laryngealization in a subsequent section.

A fifth state of the glottis which clearly contrasts with both voice and voicelessness is that during a glottal closure, when the vocal cords are held tightly together throughout their length. From a phonological point of view it is often convenient to consider a glottal stop along with articulatory stops such as p , t , k . But from a phonetic point of view it has to be considered as a state of the glottis, because of the combinatory restrictions; if there is a glottal closure there cannot simultaneously be voice, or voicelessness, or murmur, or laryngealization. Examples of contrasts involving glottal stops in Tagalog are given in Table 9.

The final state of the glottis to be discussed is that associated with whisper, in which the vocal cords are narrowed or even together anteriorly, leaving a somewhat wider gap at the other end between the arytenoid

Table 9: Contrasts involving glottal stops in Tagalog

ʔa:naj	termite	ha:naj	row
kaʔo:n	fetch	kaho:n	box
ba:taʔ	child	ba:tah	bathrobe
magʔalis	to remove	magalis	full of sores

Table 10: Contrasts between oral and nasal vowels, and examples of (allophonically) nasalized semivowels and other approximants in Yoruba

fì	use			su	scatter seed
ìjafĩ	wife			sũ	push
		ɔ̀bɔ	monkey		
		ìbɔ̃	gun		
ìjɛ̃	that	w̃ɔ̃́	they	mɛ́r̃ĩ	four

Table 11: Contrasts involving prenasalized stops in Tiv.

áa mbè	she suckled	á bèndè	he touched	á mènà	he swallowed
á ndèrà	he began	á dè	he left alone	á nèndà	he is backward in growth
á ndzùùr	he muddled	á dzèndà	he prohibited	á ɲàndè	he urinated
á ndʒɔ̀ɣɔ̀l	he spoke quickly	á dʒìŋgè	he searched		
á ŋgɔ̀hɔ̀r	he received	á gèmà	he turned round		
á ŋ͡mg͡bahom	he approached	á g͡bèr̀	he slashed		

Table 11a: Contrasts involving oral, lightly nasalized and heavily nasalized vowels in Chinantec (of Merrifield 1963). In each case the tone is a high to mid glide.

hɑ	so, such	hɑ̃	(he) spreads open	hɑ̃̃	foam, forth

cartilages. The cords are held quite stiffly, and sometimes there are additional constrictions just above the glottis. This state is linguistically significant only in situations parallel to those in which voiced sounds may be opposed to voiceless sounds. In Wolof it marks the contrast in final position between stops which, when they become non-final, are contrastively voiced as opposed to voiceless; and it is often a prosody associated with the otherwise voiced sounds in final syllables in many languages. Doke (1931) reports that it is common in the Bantu family; and it is equally typical of some forms of French.

Some other contrasts seem to me to be due to differences in the mode of vibration of the vocal cords. There appear to be two different modes underlying the contrasts between the Javanese and Indonesian sounds written with *p* and *b̄* in words such as *pipi (cheek)* and *bibi (aunt)*. The vocal cords vibrate during the same part of the closure in each case, but there is some difference in laryngeal function which is most noticeable during the release of the stop and the first part of the vowel. Different modes of vibration may also occur during the release of Korean stops (cf. Kim 1966).

The Oro-Nasal Process

The oro-nasal process is the simplest of the four major components of the speech mechanism. The soft palate, or velum, may be raised forming a velic closure in the upper pharynx; or it may be lowered allowing air to pass out through the nose. These two possibilities may be distinguished by calling the accompanying sounds either oral or nasal. The usual practice, however, is to use the term nasal for a sound in which the oral passage is blocked and all the air passes out through the nose, and the term nasalized for a sound in which the velum is lowered but there is no oral stop closure, so that some of the air passes out through the nose and some through the mouth. This obviates the necessity of calling sounds such as m, n nasal stops, and b, d oral stops. It is a convenient usage which we will follow in this section; but, as we shall see later, it may be an unmotivated multiplication of terms which should be avoided in a theory of phonetics.

Many languages distinguish between oral vowels and nasalized vowels. Nasalized semivowels, fricatives, and laterals are also quite common, but in all the languages I have investigated these sounds occur only in syllables where the vowel is also nasalized; I do not know of any contrasts between nasalized and non-nasalized semivowels in which the adjacent vowel is not similarly specified by the oro-nasal process. This is, of course, another way of saying that the oro-nasal process often affects a syllable as a whole. Examples of Yoruba contrasts of this kind are given in Table 10. Another use of the oro-nasal process common in African languages is in the formation of a series of voiced stops which contrast with other fully voiced stops by having a short nasal section during the first part of the articulation. Tiv examples are given in Table 11.

There have been claims (Peterson and Shoup 1966) that it is necessary to recognize two degrees of velic opening, but I have not heard any linguistic contrasts of this kind myself. It is true that there are allophonic variations: vowels are often nasalized to a greater or lesser extent in accordance with the nature of the adjacent consonants; and there is usually a variation in the degree of velic opening in accordance with the height of the vowel (high vowels are far less nasalized than low vowels). But these variations are all easily explicable in terms of the actions of the muscles involved. In fact Moll (1966) has shown that nearly all the well known variations in the degree of velic opening in English (and probably in other languages) can be predicted from a simple model which take into account only the accompanying articulatory gestures and an on or off instruction to the muscles responsible for the velopharyngeal closure. (Since writing this I have heard some Chinantec (Merrifield 1963). In this language there are clear contrasts, instrumentally verified, between oral, lightly nasalized, and heavily nasalized vowels, as shown in Table 11a.)

The Airstream Process

The principal airstream mechanisms are summarized in Table 12. Nearly all speech is formed with a pulmonic egressive airstream. The p physiological mechanisms involved in producing this airstream are discussed in some detail in Ladefoged (1967). Some communities use an ingressive pulmonic airstream for paralinguistic purposes, such as disguising the voice (Conklin 1949). But I have not heard of any language making a linguistic contrast by using an ingressive as opposed to an egressive pulmonic airstream. It would probably be too inconvenient (if not impossible) to reverse the action of the respiratory mechanism with sufficient dexterity.

Variations in the way in which the pulmonic egressive airstream is produced can be correlated with increases in subglottal pressure which produce significant variations in pitch (as discussed above), and with variations in stress. Phonetic literature is full of vague remarks about the nature of stress; but the data summarized in Ladefoged (1967) show quite conclusively that stress involves a gesture of the respiratory muscles, and that it can be specified in terms of the amount of work done by the air in the lungs.

Variations in the degree of activity of the respiratory muscles may also be responsible for some of the differences between fortis and lenis sounds. I doubt whether such variations actually occur during most of the sounds which have been described as being fortis as opposed to lenis. Lisker and Abramson (1964) are probably correct in saying that nearly all these sounds vary only in the relations between the time at which voicing commences and the time at which the stricture is released, and so can be specified adequately by means of terms such as voiced, voiceless, aspirated and unaspirated. But there are a few cases which cannot be explained in this way. Recent pressure recording indicate

Table 12: The principal airstream mechanisms

Airstream	Direction	Brief description	Type of stop consonant	Example symbols
Pulmonic	egressive	lung air pushed out under the control of the respiratory muscles	plosive	p t k
Glottalic	egressive	pharynx air compressed by the upward movement of the closed glottis	ejective	p' t' k'
Glottalic	ingressive	downward movement of the vibrating glottis	implosive	b d g
Velaric	ingressive	mouth air rarefied by the backward and downward movement of the tongue	click	ʇ ʗ ʖ

Table 13: Contrasts involving 'strong' (or 'double') consonants in Luganda. These fortis stops are here transcribed with [*].

paapaala	run about madly	p*aapaali	pawpaw
teeka	put	t*eeka	rule, law
kula	grow up	k*ula	treasure
bano	these people	b*ano	massive object
gaali	they were	g*aali	bicycle

Table 14: Contrasts involving lenis unaspirated stops, fortis unaspirated stops, and aspirated stops in Korean (cf. Kim 1966). The fortis stops are here transcribed with [*].

pul	fire	p*ul	horn	p^hul	grass
tal	moon, month	t*al	daughter	t^hal	mask
kətta	to walk	k*ətta	extinguished	k^hətta	grew

that the so-called strong consonants in Luganda are not only usually longer but also pronounced with greater pulmonic pressure than occurs in their weak counterparts; examples are given in Table 13. The so-called fortis stops in Korean (examples in Table 14; cf. Kim, 1966) may also be associated with greater action of the respiratory muscles, but in this language there are additional differences in the phonation process which are not present in Luganda. There are similar additional complications in the Sino-Tibetan language Jinpho', but there is clear evidence (hitherto unpublished) that the so-called fortis nasals are in this language,accompanied by a large increase in subglottal pressure.

Although the activity of the respiratory muscles can be varied in degree, the pulmonic egressive airstream mechanism cannot be turned on and off very rapidly. Consequently all sounds in all languages are produced while the air in the lungs is at above atmospheric pressure. But there may also be supplementary airstream mechanisms involved. Movements of the vocal cords are used in the glottalic airstream mechanism to produce the air pressure variations in the two types of glottalic sounds, ejectives and implosives. Ejectives are formed by bringing the vocal cords tightly together and then raising and constricting the whole larynx, so that the pressure of the air in the mouth and pharynx tends to be raised. Contrasting series of stop consonants are made in this way in many African, Amerindian, Caucasian and other languages, some of which also have ejective fricatives. Examples of both stops and fricatives in Amharic are given in Table 15.

It is perfectly possible to produce ingressive glottalic sounds by a similar process in which the closed glottis is rapidly lowered instead of raised. Stops using this mechanism have been reported in an Amerindian language, Tojolabal (Pike, E., 1963), but I have not heard them myself. The more common airstream process involving the lowering of the glottis does not have the vocal cords held tightly together. Instead, as they descend they are allowed to be set in vibration by the air in the lungs, which is always at a higher than atmospheric pressure during any speech activity. The action of the vocal cords in the production of these implosive sounds has been described as that of a leaky piston. Often the piston is so leaky that the airstream is not actually ingressive nor the sounds really implosive. In many of the languages I have observed (cf. Ladefoged 1964) the pressure of the air in the mouth during an ingressive glottalic stop is approximately the same as that outside the mouth, since the rarefying action of the downward movement of the glottis is almost exactly counterbalanced by the leakage of lung air up through the vocal cords. Although these sounds may be called implosive, in ordinary conversational utterances air seldom flows into the mouth when the stop closure is released.

Many of the languages which have ejectives also have implosives, but usually (perhaps always?) at different points of articulation. Among other languages with implosives, Sindhi also has voiced, voiceless

Table 15: Contrasts involving ejective stops and fricatives in Amharic.

t'ɨl	quarrel	tɨl	warm	dɨl	victory
k'ɨr	stay away	kɨrr	thread	gərr	innocent
mətʃ'	one who comes	mətʃ	when	mədʒ	grinding stone
s'əgga	grace	səgga	to worry	zəgga	to close

Table 16: Contrasting stops in Sindhi

	ɓəni	curse	bənu	forest	pənu	leaf	pʰəṇu	snake hood	b̤ənəṇ̤u	lamentation
			dəru	door	təru	bottom	tʰəru	district name	d̤əru	trunk of body
	kʰəɗo	pit	gəḍo	dull	kʰəṭo	sour	kəṭʰo	assembled	kəḍ̤o	take out
*	bəʄu	run	ɟəɟu	judge	bəcu	be safe	bəcʰu	attack	vəɟ̤u	opportunity
	səɠi	braid tail	bəgi	buggy	ʃəki	suspicious	səkʰi	girl friend	səg̤i	healthy

* ɟ and c are affricates, and might be transcribed <u>dʑ</u> and <u>tɕ</u> .

Table 17: Contrasting clicks in Zulu. All these items are imperative forms of verbs, all with the tone pattern low - high.

	(laminal) dental		(apical) alveolar		alveolar lateral	
aspirated	ʇʰaʇʰa	be evident	ʗʰaʗʰa	rip open	ʖʰoʖʰa	stab, jab
voiceless unaspirated	ʇaʇa	climb	ʗaʗa	explain	ʖoʖa	narrate
voiced	g͡ʇag͡ʇa	dance at wedding	g͡ʗoka	dress up	g͡ʖoba	pound
murmured nasal	ŋ̤͡ʇoŋ̤͡ʇa	gather unripe corn	ŋ̤͡ʗala	tie tightly	ŋ̤͡ʖaŋ̤͡ʖa	coax
voiced nasal	ŋ͡ʇaŋ͡ʇa	act quickly	ŋ͡ʗoŋ͡ʗa	resound	ŋ͡ʖola	slope

unaspirated, aspirated and murmured stops, so that there are 24 contrasting stops in an almost complete five by five array as shown in Table 16. (The situation is complicated by the fact that the so-called palatal explosives are actually pre-palatal affricates; but the voiced implosive is a true palatal stop.)

Some languages, such as Swahili and Marwari, have implosives as free variants or allophones of voiced pulmonic stops (or plosives). The difference between implosives and plosives is one of degree rather than of kind. In the formation of voiced plosives in many languages (e.g. English, cf. Hudgins and Stetson, 1935) there is often a small downward movement of the vibrating vocal cords so as to allow a greater amount of air to pass up through the glottis before the pressure of the air in the mouth has increased so much that there is insufficient difference in pressure from below to above the vocal cords to cause them to vibrate. An implosive is simply a sound in which this downward movement is comparatively large and rapid.

There are a number of other cases in which the categories we have been defining are not completely discrete. Downward movements of the larynx in some languages are often accompanied by a tendency towards laryngealization. These sounds are in some senses both implosives and laryngealized stops; and in fact no language uses the difference between these two possibilities. Similarly there are no linguistic contrasts and it is impossible to distinguish absolutely between a laryngealized stop, in which each opening of the vocal cords is separated by a fairly long closed period, from a plosive with regular voicing which is interrupted by a comparatively short glottal stop. There are also no recorded cases of a language using a contrast between an ejective and glottal stop. (But some languages, such as Huixteco Tzotzil, have contrasts between an ejective affricate and an affricate followed by a glottal stop.)

There are other difficulties in that the categories we have outlined so far may need extending to cope with phenomena that have been observed in some languages. I am not altogether certain about the Jinpho' nasals which were mentioned earlier, partly because there seems to be a parallel phenomenon in the voiceless stops. Jinpho' has three series of stops, one which is clearly voiced, one which is clearly voiceless and aspirated, and one, auditorily in between the other two, which is largely voiceless and quite unaspirated in which the vocal cords start vibrating and the glottis starts descending shortly before the release of the closure. We do not have a category which will specify what is in common between the (voiced) nasals and these (voiceless) stops, unless we attach an arbitrary label such as tense or fortis to both; but, although it might be phonologically convenient, how we would define the phonetic properties of this term I do not know.

It is perhaps worth noting that the term glottalized has been avoided in all the preceding discussion, largely because it has been used by others in so many different ways. It might be appropriate as a cover term for ejectives, implosives, laryngealized sounds, and pulmonic articulations accompanied by glottal stops. But it is otherwise not very useful in precise phonetic descriptions.

There is still another airstream mechanism which we have not yet discussed. This is the velaric airstream mechanism in which a body of air is enclosed by raising the back of the tongue to make contact with the soft palate, and either closing the lips or (more commonly) forming a closure on the teeth or alveolar ridge with the tip (or blade) and sides of the tongue. The air in this chamber is rarefied by the downward and backward movement of the body of the tongue, the back of the tongue maintaining contact with the soft palate. When a more forward part of the closure is released, air rushes into the mouth, and a sound known as a click is produced. This mechanism is always ingressive, and there are no reports of its use in the formation of sounds other than stops, although it is perfectly possible to produce velaric ingressive fricatives.

The Khoisan languages (Bushman and Hottentot) and some of the neighbouring Southern Bantu languages such as Xhosa and Zulu are the only groups using a velaric airstream mechanism to produce sounds made at a number of different places of articulation. But the air in the mouth is sometimes rarefied in a similar way during the production of labial velar stops in West African languages (cf. Ladefoged 1964). Zulu and Xhosa have three sets of clicks, dental, alveolar lateral, and retroflex; the first two of these often have a slightly affricated quality. The Khoisan languages have clicks at other places of articulation. Lanham (1964) says that the different languages in this group commonly use four out of the six possibilities: bilabial, dental, alveolar, alveolar lateral, palatal, and retroflex. No one language is reported to have more than five of them. Speakers of the Khoisan languages are no longer readily available, and I have not been able to hear any of them myself.

Since the velaric airstream mechanism involves only movement of mouth air and a velar closure, the pulmonic airstream which is always present in speech can be used to produce a velar plosive or a velar nasal which may be formed at the same time as the click, and released slightly afterwards. In addition it is possible for any of the different phonation types to be used. Zulu (like other Nguni languages) has voiced, voiceless unaspirated, and aspirated clicks (velaric sounds accompanied by g , k , and k^h); and voiced nasal and murmured nasal clicks (velaric sounds accompanied by ŋ and ŋ̤) as shown in Table 17. These clicks parallel contrasts between other sounds in the language; Zulu also has three way contrasts among pulmonic stops, and a two way contrast between voiced and murmured nasals.

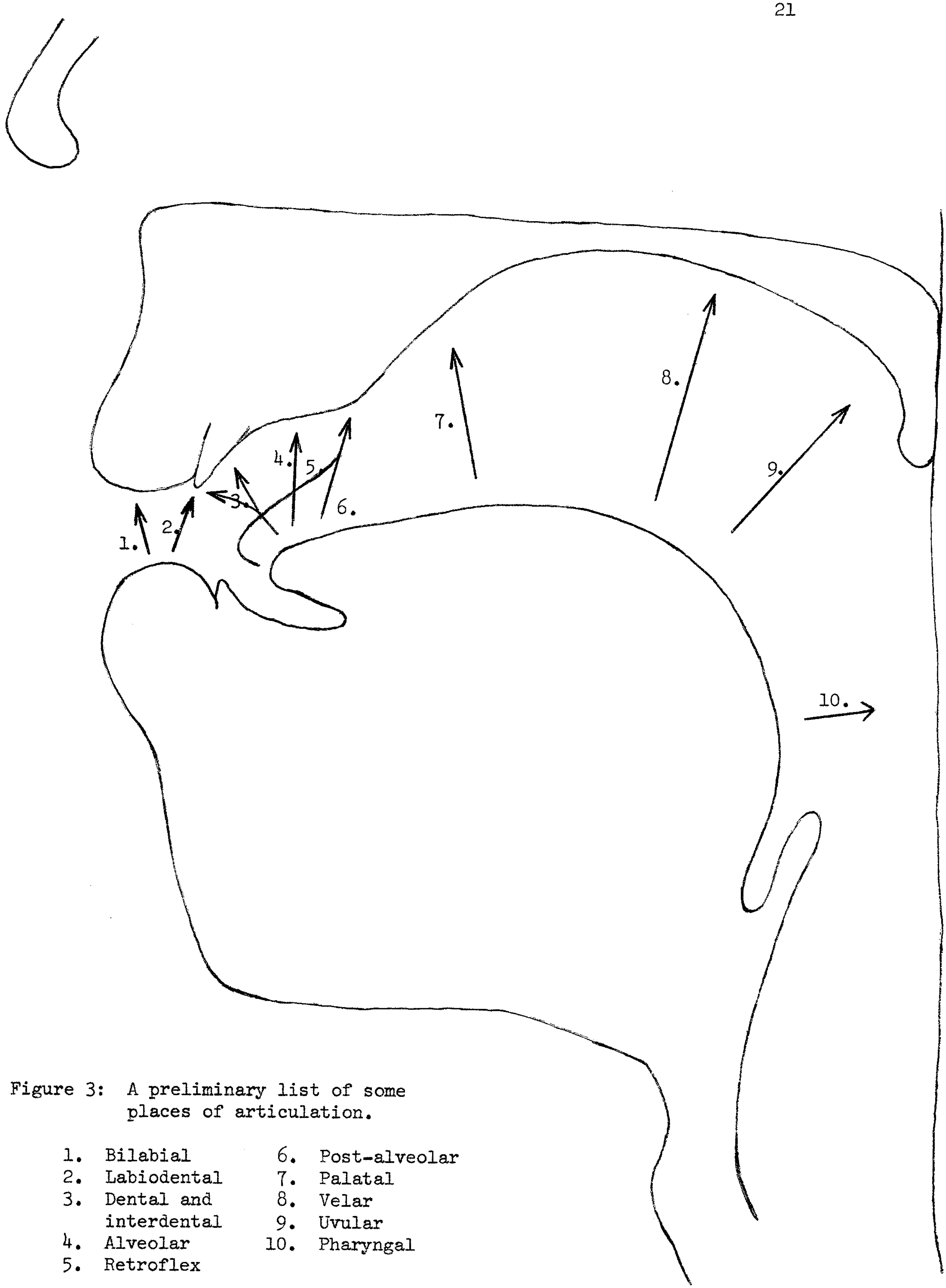


Figure 3: A preliminary list of some places of articulation.

1. Bilabial
2. Labiodental
3. Dental and interdental
4. Alveolar
5. Retroflex
6. Post-alveolar
7. Palatal
8. Velar
9. Uvular
10. Pharyngal

The Articulatory Process

The articulatory process is the most complicated of all the four processes. It should probably be considered as consisting of two or three independent sub-processes. We shall do this in that we will follow the usual approach of considering the actions of the tongue and lips which make up this process first in terms of the possible places, and then in terms of the possible manners of articulation.

Places of Articulation

One of the difficulties of talking about places of articulation is that neither the tongue nor the roof of the mouth can be divided into discrete sections. The teeth are arranged on the alveolar ridge in such a way that it is difficult to divide dental from alveolar. The end of the alveolar ridge has been defined by Jones (1956) as the point where the roof of the mouth ceases being concave and becomes convex; but when one looks at actual data on mouth shapes it is often difficult or impossible to locate this point; and in any case it makes the center of the alveolar region much further back than in the articulations phoneticians commonly call alveolar--the diagrams of English alveolar t in Jones's works show an articulation more retracted than that made by any English speaker I have investigated. The center of the palatal region is equally hard to define, although here Jones has a more useful concept in his use (e.g. in Jones 1956, and in the I.P.A. *Principles*, 1949) of the term cardinal palatal to mean (apparently, I cannot find an explicit definition to this effect) an articulation in the area of cardinal vowel number one (the highest and most front vowel possible). The category velar can be defined by relation to the soft palate as opposed to the hard, but this division is also difficult to locate in practice. And the distinction between velar and uvular is even more arbitrary. In fact, when discussing places of articulation it seems especially necessary to bear in mind that the categories are required simply for distinguishing linguistic oppositions.

A preliminary set of categories of place of articulation is represented diagrammatically in Figure 3. Each of the terms defines an action of both a lower and an upper articulator. There is little difficulty about the first two terms. Bilabial stops and nasals occur in nearly all languages. Some languages also have bilabial fricatives which have to be distinguished from labiodental fricatives; examples from Ewe which contrasts both voiced and voiceless sounds of this kind are shown in Table 18. From my own work I know of contrasts between bilabials and labiodentals only among fricatives and what we shall call approximants (frictionless continuants and semivowels). Labiodental nasals occur in many languages, but probably only as intrinsic allophones of other nasals. Labiodental stops contrasting with bilabial stops have been reported in Tonga (Guthrie 1948), but I did not succeed in eliciting them from my Tonga informants; I have, however, heard a labiodental stop made by a Shubi speaker with good teeth, but only as an allophone of pf . Nor have I heard a lingualabial stop, in which

Table 18: Contrasts involving bilabial and labiodental fricatives in Ewe.

ɛ̀βɛ̀	the Ewe language	ɛ̀vɛ̀	two
éɸá	he polished	éfá	he was cold
èβló	mushroom	évl̩ó	he is evil
éɸlè	he bought	éflé̂	he split off

Table 19: Symbols and terminology for dental and alveolar articulations, and examples from West African languages; the laminal articulations are usually affricated.

Active articulator	tip	tip + blade	tip	blade	tip	blade
Passive articulator	teeth	teeth + teeth ridge	front of teeth ridge		back of teeth ridge	
Label summarizing active and passive articulators	(apical) dental	(laminal) denti-alveolar	apical alveolar	laminal alveolar	(apical) post-alveolar	(laminal) pre-palatal
Example of symbol	d̪	d͔	d	d͕	ḍ	d̠
Exemplifying language						
Temne	t̪or descend			t͕or farms		
Isoko		òt͔ú louse	òtú gang			
Ewe		éd͔à he throws			éḍà she laid	
Twi			ɔ́dà he lies			ɛ̀d̠á father

the tongue makes contact with the upper lip; this sound is mentioned by Hockett (1955) as occurring in Bororo, but I gather from Lounsbury (personal communication) that the reference should really have been to another South American language, namely Umotina.

It may be necessary to distinguish between dentals and interdentals, but I do not know of any use of this distinction. However the categories dental, alveolar and post-alveolar undoubtedly need further differentiation. Sounds in each of these three categories can be made with either the tongue tip (in which case they are called apical) or the tongue blade (in which case they are laminal). There are thus six possible combinations; but no language uses more than three of them. Some of the Dravidian languages, such as Malayalam, have three stops and three nasals in this region, all of them being made with the tip of the tongue. The Dravidian languages are the only ones I know in which contrasts involving these nearby places of articulation on the roof of the mouth do not also involve a different part of the tongue. Among West African languages the contrasts are between apical dentals and laminal alveolars, or between apical alveolars and laminal dentals (in which the tongue contact usually extends beyond the teeth, so that they are really denti-alveolars). Table 19 lists some languages using these sounds, together with exemplifying symbols and a possible terminology which can be used if it is thought desirable to incorporate the apical - laminal distinction into the set of terms applying to place of articulation. Examples from these West African languages are given in Table 19; these contrasts have been illustrated by instrumental data and discussed in detail elsewhere (Ladefoged 1964). The contrasting voiceless stops and nasals in Malayalam are illustrated in Table 20; the bilabial, palatal, and velar nasal articulations are also given, so that it may be plainly seen that this language has six contrasting places of articulation.

Both apical post-alveolar (retroflex) sounds and laminal post-alveolar (prepalatal) sounds cover a wide range of articulatory areas. In some Indian languages the retroflex consonants involve only the tip of the tongue and the back of the alveolar ridge, whereas in others there is contact between a large part of the under side of the tongue tip and much of the forward part of the hard palate. Similarly the actual area of contact in prepalatals may vary over quite a wide range, so that it is often hard to decide whether a given sound should be classified as a prepalatal or a palatal. In fact, it seems probable that no language distinguishes between sounds simply through one being a palatal and the other a prepalatal. All the languages which use these two articulatory positions (such as Ngwo, illustrated by palatograms in Ladefoged, 1964) either have affricates in the one position and stops in the other, or in some other way supplement the contrasts in place of articulation with additional variations in the manner of articulation. We may have no justification for attempting to distinguish between these two categories.

Table 20: Contrasts involving bilabial, dental, alveolar, post-alveolar, palatal and velar stops and nasals in Malayalam, illustrating the necessity for six points of articulation.

Bilabial	Dental	Alveolar	Post-alveolar	Palatal	Velar
	mut̪t̪u pearl	muttu density	muṭṭu knee		
	kut̪t̪i stabbed	kutti peg	kuṭṭi child		
kʌmmi shortage		kʌnni Virgo	kʌṇṇi link in chain	kʌɲɲi boiled rice and water	kʌŋŋi crushed
	en̪n̪ʌ named	ennʌ me	eṇṇʌ oil	te:ɲɲʌ worn out	te:ŋŋʌ cocoanut

Table 21: Contrasts involving palatal, velar and uvular stops in Quechua. (The palatal stops are strongly affricated.)

ciri	cold	cʰilci	drizzle	c'ici	dirty
kiru	tooth	kʰipu	*Kipu	k'icij	pinch
qara	skin	qʰapan	rich	q'acu	grass

Table 22: Contrasts involving voiced and voiceless pharyngeal approximants or fricatives, murmured approximants, and glottal stops in Arabic.

ʕala	on	ʕamm	paternal uncle	waaʕʕa	container
ħallaʔ	right away	ħamm	concern		
ɦalaʔ	shave	ɦammaam	bath	waaɦed	one
ʔaḷḷa	God	ʔamar	command	waʔʔaf	stop

Contrasts between palatals, velars, and uvulars are illustrated in Table 21, which lists examples from Quechua. The palatal stops in many languages tend to be more affricated than the others, perhaps because of the mechanical difficulty of quickly withdrawing the front of the tongue, which often contacts a large area of the roof of the mouth in the formation of these stops. I do not know of any language which distinguishes between a palatal stop and either a palatal affricate or the sequence palatal stop followed by palatal fricative.

Stops, nasals, and fricatives occur at all the places of articulation discussed so far, with the exception of the labiodental category. The illustrations given have been mainly of stops and nasals; fricatives and other manners of articulation such as laterals and trills will be considered later. In the pharyngal area, however, no language uses stops (most people cannot make them), and nasals are an impossibility. Even fricatives are not very common. Table 22 gives examples (suggested by Ferguson, 1961) of Arabic fricatives and some auditorily similar sounds with different phonation types in which (in the terminology being developed here) there is no specific place of articulation.

In addition to the places of articulation we have discussed so far there are also three double articulations, each of which involves the actions of both lips and, simultaneously, some part of the tongue and the roof of the mouth. We shall use the terms labial-velar, labial-palatal, and labial-alveolar when describing these sounds, restricting the form labio- to the term labiodental, which specifies an articulation involving only one lip. In a double articulation the two lips and the other articulators are both involved to an equal extent in the production of the sound. These articulations should be distinguished from those in which there is a primary articulation at some point and additional lip-rounding forming a secondary articulation, so that the sound may be said to be labialized. Labialization and other secondary articulations will be discussed later.

The approximant (semivowel) w in English and many other languages is the most familiar labial velar sound. Labial velar stops, nasals, and fricatives also occur in West African languages, as I have described in detail elsewhere (Ladefoged 1964). The stops and nasals are usually symbolized by the bilabial and velar symbols joined together by a tie bar, to indicate that the articulations are simultaneous and not sequential. Examples of contrasts in Idoma are given in Table 23. When there is a contrast between a labial-velar approximant and a labial-velar fricative it may be symbolized by adding the diacritic which indicates a closer articulation to the symbol for the approximant. Margi examples of this kind (and of the contrast between a palatal approximant and a palatal fricative) are given in Table 24. The contrast between a labial-velar and a labial-palatal approximant occurs in French (examples in Table 25) Idoma, Kamba, etc. I do not know of any examples of labial-palatal stops or nasals, but the Shona 'whistling fricatives' illustrated in Table 26 may exemplify another

Table 23: Contrasts involving labial-velar, labial, and velar stops and nasals in Idoma.

àk͡pà	bridge	àpà	lizard	àka	wheel
àg͡bà	jaw	àbà	palm nut	àga	axe
aŋ͡màa	painted body marks	áma	bell	aŋá i	Western rainbow

Table 24: Contrasts involving palatal approximants and fricatives, and labial-velar approximants and fricatives in Margi.

jà	give birth	j^àj^àd̤ò	picked up	káwà	sorry	w^á	reach inside

Table 25: Contrasts involving palatal, labial-palatal, and labial-velar approximants in French.

Palatal		Labial-palatal		Labial-velar	
mjɛt	crumb	mɥɛt	mute	mwɛt	sea gull
lje	tied	lɥi	him	lwi	Louis
		ɥit	eight	wi	yes

Table 26: Contrasts involving the 'whistling fricatives' (here symbolized ɕ̣w ʑ̣w) in Shona. In going from s to ɕ͡w to ɕ there is increasing retardation and flattening of the tongue.

(apical) alveolar		labial-(laminal) alveolar		(laminal) pre-palatal	
màsóró	big heads	ɕ͡ẉòɕ͡ẉé	sugar ant	mùɕòmà	be hoarse
màzòrò	turns	ʑ͡ẉósé	all	ʑòʑómá	tuft of hair
tsàmà	handful	ìtɕ͡ẉá	new thing	t̠ɕàná	fat child
dzàmá	disappear	ìdʑ͡ẉá	new thing	d̠ʑànà	turn

manner of sound made with this double place of articulation; these sounds may, however, be more properly called labial-alveolars. The only other labial-alveolars I have heard are stops or nasals. They occur in West African languages, often (as in Dagbani) as allophonically predictable variants of the more common labial-velars, but in some languages (such as Bura, illustrated in Table 27) as the only series of double articulations. I have not heard any contrasts between labial-alveolars and labial-velars, but they are reported by Chinebush (1965) in Nzima.

Many other double articulations are possible. An alveolar-velar fricative has been reported in some dialects of Swedish (Abercrombie 1967); and I have heard what might be labiodental-postalveolar fricatives in Kutep (illustrated by X-ray data in Ladefoged, 1964). But I am not sure if these sounds have two equal articulations, each producing a turbulent airstream. There are more obviously two articulations in the Shona nasals written *ny*. Both my own palatograms and those of Doke (1931) show that there are two contacts--a tongue tip (or blade) and alveolar ridge contact, and probably simultaneously, a tongue front and hard palate contact. But these contacts are probably best considered to be due to accidents of the shapes involved, rather than deliberate double articulations of the kind discussed above.

Manners of Articulation

A preliminary list of the manners of articulation is given in Table 28. As a matter of convenience in summarizing terms the two possible states of the oro-nasal process are also listed in this table. This list of manners of articulation is clearly insufficient to account for all the linguistic contrasts that occur. But it is useful in that it allows us to begin with a simple model which distinguishes between only three degrees of obstruction to the airstream (complete stoppage, restricted flow, and unimpeded flow; here called stop, fricative, and approximant) and three other gestures (trill, tap, and flap), which occur at a more limited number of places of articulation. For the moment we may consider the approximant category to be simply a convenient general term to include what others have called semivowels, laterals and frictionless continuants (as well as vowels, which we will consider later).

Within each of the groups of terms in Table 28, specification of any one term precludes all the others. If a sound is oral it is not also nasal; if it is a stop, it cannot be simultaneously a fricative, or an approximant, or a trill, or a tap, or a flap; if it is central it cannot be lateral. But between terms in different groups there is a great deal of freedom of combination. All these articulations can be oral or nasal (although, as we have already noted, possibly only stops and approximants contrast in this way at the segmental level); and we shall see that the terms central and lateral can be applied not only to approximants and fricatives, but also to stops and some forms of flaps.

Table 27: Contrasts involving labials and labial-alveolars in Bura.

pàkà	search	p͡tá	hare
psá	lay eggs	p͡tsà	roast
pʃàrì	spread a net	p͡tʃi	sun
b̰aɬa	dance	b̰͡d̰à	chew

Table 28: The oro-nasal process and some articulatory categories

Phonetic term	Brief description	Example symbols
Oral	soft palate raised forming a velic closure; none of the air going out through the nose	a w b
Nasal	soft palate lowered so that some or all of the air goes out through the nose	ã w̃ m ŋ͡ɬ
Stop	complete closure of two articulators	p b m ʗ
Fricative	narrowing of two articulators so as to produce a turbulent airstream	v s ɬ
Approximant	approximation of two articulators without producing a turbulent airstream	a w ẉ
Trill	one articulator vibrating near another	r ʀ
Tap	one articulator thrown against another	ɾ
Flap	one articulator striking another in passing	ṛ ɹ
Lateral	articulated so that air passes out at the side	l ɬ ɺ ʗ
Central	articulated so that air passes out in the center	w s ɹ ʇ

Many of these combinations are abbreviated or left not fully specified in traditional phonetic terminology. We have already noted that oral stops and nasal stops are more commonly called simply stops and nasals. Similarly central fricatives and central approximants are usually referred to simply as fricatives and approximants, and lateral approximants are usually called laterals. Generally, combinations are noted only in describing less common phenomena such as lateral fricatives.

Stops and nasals have already been plentifully exemplified, and need not be considered further. But the category fricative requires more discussion for a number of reasons. In the first place it should be noted that the turbulence of the airstream is not necessarily formed at the actual place of articulation--the point at which the two articulators are closest together. The two places coincide in the formation of v ; but in s the principal source of acoustic energy is the turbulence produced when the jet of air, which is formed by the groove between the tongue and the alveolar ridge, strikes the edges of the teeth. Another difference among fricatives is that some are made with the tongue relatively flat in the mouth, whereas others involve the formation of a comparatively narrow groove. Some writers have recognized this by further division of the members of the class which are made with the fore part of the tongue, separating out a class of grooved fricatives (Pike 1943), or rill spirants (Hockett 1955). It is often difficult to decide how to apply this distinction; and it is clearly irrelevant in the case of fricatives made with the lips and the back of the tongue. Nevertheless there are good reasons for thinking that some further division of the fricative category is necessary. A palatographic investigation (Ladefoged 1957) showed that in some people's speech the sounds at the beginnings of the English words *sip* and *ship* were articulated on very similar parts of the alveolar ridge; the consistent difference in these two sounds for all English speakers was not in the place of articulation, but in the fact that ʃ was always associated with a wider articulatory channel and more doming (as opposed to hollowing) of the fore part of the tongue. Using only the categories we have established so far; the only way to differentiate between these two sounds, both of which may be made with the tip of the tongue, is by calling one apical alveolar and the other apical postalveolar. But if we do this we cannot differentiate between an apical ʃ and the apical postalveolar which we previously called retroflex. These two sounds may contrast in many Dravidian languages, all of which also have an apical s which is not particularly dental. Kannada and Telugu examples are given in Table 29. I have heard a number of Kannada informants who have two sounds almost identical with those in English *sip* and *ship* (except that they do not have the accompanying movements of the lips which normally occur in English) as well as a third sound which is a voiceless retroflex fricative with (to me, but not to them) a very similar auditory quality. But the situation is complicated because the difference between the apical ʃ and the apical postalveolar fricative in these Dravidian languages is associated with erudition and

Table 29: Contrasts involving apical alveolar, postalveolar, and retroflex fricatives in Kannada, and in Telugu.

	alveolar	postalveolar	retroflex
Kannada	saaku enough	ʃaama black	wiṣa poison
Telugu	maasaw month	aaʃa ambition	kaṣaajaw decoction
	haasjʌw̃ sarcasm	druʃjam scenery	b̤aasjʌw̃ commentary

Table 30: Contrasts involving alveolar taps and trills in Tamil and Spanish.

	alveolar tap	alveolar trill	postalveolar approximant
Tamil	əɾəm saw	ərəm charity	aaɹəm depth
Spanish	peɾo but	pero dog	

Table 31: Contrasts involving two different apical trills, a retroflex approximant and an alveolar and a retroflex lateral in Malayalam.

advanced alveolar trill	retracted alveolar trill	retroflex approximant	alveolar lateral	retroflex lateral
ʌr˂ʌ half	ʌr˃ʌ room	aaẓʌ banana tree		
pur˂ʌ roof	pur˃ outer	puẓʌ river		
kʌr˂i charcoal	kʌr˃i curry	kʌẓi skein of yarn	kʌli possessed by a spirit	kʌḷi game

Table 32: Contrast involving a flap and a tap (which may be allophonically a trill) in Hausa.

báɽà servant — báɾà begging

a spelling difference. Many people who claim to make the difference do not always do so in their everyday speech.

One way of differentiating between all these fricatives would be to list an additional place of articulation, and regard (apical) postalveolar and (apical) retroflex as distinct categories. But this seems unsatisfactory since this extra place is needed only to account for the fricative sounds which occur, and we have already noted that it is not really the appropriate way of distinguishing among these. Another solution would be to specify the difference between s and ʃ in terms of the secondary articulation of palatalization (to be described later). But this does not account for the differences in the articulatory channel described above. At the moment I am inclined to think that the best solution is to distinguish between grooved and slit fricatives in this area. But perhaps it would be better to admit that we do not have a good way of describing these tongue tip articulations, especially when we come to examine some of the other problems, such as the difference between ʒ , z and ɹ a voiced apical fricative which is nevertheless not retroflex. These three sounds occur in contrast in some languages (e.g. South African English where ɹ is the usual form of r). In a previous discussion (Ladefoged 1964) of Bini, which also has ɹ , I carefully avoided the necessity of specifying this sound in terms of exact categories.

A fricative form of trilled r occurs in Czech. This leads us to consider whether the category fricative is really a member of a set of which the other members are stop, approximant (or something equivalent), trill, tap, and flap, or whether it is, like central - lateral, an independent, additive component. What characterizes the Czech variant of the trill manner of articulation is that it is a laminal (and not an apical) trill; so this example does not provide us with good grounds for saying that the category fricative is not a member of the larger set, mutually exclusive with all the other terms. In any case, no language distinguishes between a fricative and a non-fricative trill, so we have no motivation for introducing this possibility into our scheme of categories. (This does not, of course, dispose of the question of whether the category fricative is an independent, additive component or not. We will return to this point later.)

The other manners of articulation are much more restricted in the places of articulation at which they can occur. The majority of trills, taps, and flaps are made with the tip of the tongue. In a trill the tip of the tongue may be loosely held near some part of the roof of the mouth and set in vibration by the action of the airstream in much the same way as the vocal cords are set in motion during the production of voice. In a typical speech sound produced in this way there may be about three vibratory movements; but even in cases where there is only a single contact with the roof of the mouth, the action is physiologically (but perhaps not auditorily) quite distinct from that of a tap. A tap is formed by a single contraction of the muscles such that one articulator is thrown against the other. The distinction between these two gestures is exemplified in Table 30 by contrasts in Tamil and Spanish. Educated Tamil also has a voiced postalveolar approximant, so that some speakers have three contrasting sounds similar to the sounds which are all diaphones of r in different forms of English.

Apical trills or taps are usually in the dental or alveolar regions. Malayalam is the only language I have investigated which (in some dialects) makes a distinction between two trills in this area, one being more dental and the other more alveolar. A recent palatographic investigation showed that these trills are probably further distinguished by the action of the back of the tongue. These two trills contrast with two laterals, and with a retroflex approximant as shown in Table 31.

A flap is an articulation which usually involves the curling of the tip of the tongue up and back and then allowing it to hit the roof of the mouth as it returns to a position behind the lower teeth. A flap is therefore distinguished from a tap by having one articulator strike against another in passing while on its way back to its rest position, as opposed to striking immediately after leaving its rest position.

Retroflex flaps are common in Indo Aryan languages, but, in these languages, they do not contrast with taps or trills. A distinction between a flap and a tap (which may be allophonically a trill) is made in Hausa. Palatograms and details of the contrast have been published elsewhere (Ladefoged 1964); examples are given in Table 32. Voiceless trills, taps, and flaps are comparatively uncommon; many languages (e.g. Gaelic, Bini) have a voiceless fricative or approximant ɹ̥ , but I have not heard any forms where r̥ or ɾ̥ is the normal variant.

The central lateral dichotomy may be applied to flaps, but not to taps and trills. There are a number of languages in which sounds having the characteristic gesture involved in making a flap may have in addition a distinctly lateral quality; when the articulation is formed there is contact only in the center of the mouth, so that momentarily there is a position similar to that of an l . This kind of sound often occurs in languages which do not make a contrast between l and any form of r (e.g. Haya); but it also occurs as a third item contrasting with both l and some form of r in a number of languages. Some Chaga examples are given in Table 33.

The tongue tip is not the only articulator which can be trilled: the uvula can be made to vibrate in the same way. No language contrasts uvular and lingual trills; nor does any languages contrast uvular trills with uvular fricatives and approximants. Different dialectal forms of

Table 33: Contrasts involving an alveolar trill, and alveolar lateral flap, and an alveolar lateral approximant in Chaga.

riha	to mash	ɺiho	exciting	litʃa	something good
rina	a hole	ɺika	hide something		

Table 34: [This table has been deleted]

Table 35: Contrasts involving a labial trill or flap in Ngwe; this articulation is here indicated by *.

m*ɤ:	tadpoles	bə̀vɛt	grease
mάfɔ̀ɑ̀	chieftainness	mbɛm	seed

Table 36: Contrasts involving labiodental flaps in Shona; this articulation is here indicated by *.

ko*ó	ideophone indicating blackness
wó*o	ideophone indicating movement

Table 37: Contrasts involving laterals in Zulu. All these items are imperative forms of verbs, all with tones low - high.

	alveolar lateral approximant	alveolar lateral fricative	alveolar lateral velaric stop (click)	palatal lateral glottalic stop (ejective)	alveolar central velaric stop (click)
voiced	londa preserve	ɮuɮa roam loose	g͡ʖoba pound		g͡ǃoka dress up
voiceless		ɬoɬa prod	ʖoʖa narrate	cʎ'ecʎ'a tattoo	ǃaǃa explain

French use all these possibilities.

Trills, taps, and flaps may also be made with the lips. Again these three possibilities are not used contrastively; but at least the first two occur in different languages. I have heard a voiced bilabial trill or flap in Ngwe, which is in phonemic contrast with m , b , mb (Dunstan 1964), but which occurs only after m . It is therefore a trill generated in a slightly different way from the tongue tip trill, in that the lips are blown apart by the pressure of the airstream, instead of being held loosely apart and then being sucked toghether and set into vibratory motion. Other lip actions occur elsewhere. Margi has a kind of labiodental flap, photographs of which have been published elsewhere (Ladefoged 1964). The articulation is fairly complex and does not fit into any of the categories defined above. The lower lip is first pulled backwards against the upper teeth, contact being maintained at the same time with the upper lip. Then there is a downward movement so that the lower lip comes away from the upper lip, slips off the upper teeth, and (because of the backward pull which is being exerted) moves in behind the upper front teeth. As it is brought back from this position to its normal position, it flaps against the upper teeth. A similar sound occurs in Shona; good photographs have been published by Doke (1931). My Shona informants pronounced this sound without the first stage, the tensing of the lower lip against the upper teeth, found in Margi; in Shona the lower lip is simply drawn back behind the upper teeth, and then flapped forward with a much looser action. Some words contain two of these flaps in succession. Similar sounds have been reported in other African languages (Westermann and Ward 1933, Tucker 1940). Bilabial trills have also been noted in the Amerindian languages Amuzgo and Isthmus Zapotec (Pike, E. 1963). But in accordance with the general scheme of this book Tables 35 and 36 list only Ngwe and Shona examples which I have heard myself. Voiceless varieties of these sounds seem to be unknown.

The central - lateral opposition is completely independent of the categories specifying manner of articulation. It can be applied to approximants (the ordinary l sound in English is an alveolar lateral approximant; Malayalam retroflex lateral approximants were illustrated in Table 31), and to flaps (as in the Chaga examples in Table 33), and fricatives (examples of Zulu alveolar lateral fricatives are given in Table 37). The contrast between lateral approximants and lateral fricatives occurs only among voiced sounds. I do not know of any language that distinguishes between voiceless lateral fricatives and approximants, although many languages (e.g. Welsh and Burmese, see Table 2) have one or other of these two sounds. As shown in Table 37, Zulu has a voiceless alveolar lateral fricative, as well as a contrast between a voiced alveolar lateral approximant and fricative. In addition, Table 37 includes some items labeled lateral stops. In this label it might appear that the term lateral is being used in an unusual way; but almost the only way of distinguishing between the sounds at the beginning of the Zulu words for *pound* and *dress up* is by calling one of them an alveolar lateral click, and the other an alveolar central click. It is theoretically possible to regard these items as sequences, but of which segments I am not at all clear. Consequently in practice it seems difficult to avoid applying the terms central and lateral to clicks; and once the practice has been established of regarding certain stop consonants (the clicks) as being central or lateral, it seems logical to extend this usage to ejectives (and, in other languages, to implosives and plosives). There is clearly a physiological unity to ejective laterals (such as Zulu cʎ̥’ and the more common tl̥’ which forms part of the ejective series in many Amerindian languages such as Navaho). It is very difficult to describe these gestures as sequences of two other gestures. So we may regard the central - lateral opposition as an additive component for stops, as well as for fricatives, approximants, and flaps. When used in relation to stops this opposition specifies not just the manner of the sound after the closure, but more especially the place of release of the closure. Although the terms central and lateral were summarized in Table 28 along with other terms such as stop, trill, and fricative, we must remember that central and lateral really form an independent set, just as much related to the terms specifying the place of articulation as to those specifying the manner.

Lateral articulations probably occur only with the dental, alveolar, retroflex, and palatal categories of articulation. Velar laterals are reported (... by whom? ...) in some Malayopolynesian languages; but I have heard only voiced velar fricatives or approximants in the languages of this group which might have this sound (e.g. Aklanon). Investigators may be tempted to imagine that this sound is a lateral simply because in neighbouring languages there is an l in cognate words.

Secondary Articulation

Sounds can also be modified by secondary articulations which occur at the same time as the primary articulations. Following Pike (1943) we will consider a secondary articulation to have a lesser degree of stricture than the primary articulation, and to be made with articulators left free by the primary articulation. The most important secondary articulations are movements towards and away from an approximant position occurring simultaneously with the formation and release of another articulation. Some secondary articulations which may be needed to account for linguistic contrasts are listed in Table 38. (It should be remembered that nasalization and phonation types such as laryngealization are not articulations, and therefore not considered here.)

Labialization, an approximation or rounding of the lips, may be exemplified as the main distinguishing feature in several pairs of sounds in Akan languages, as in the Twi examples shown in Table 39. It is also a component of the pronunciation of r, ʃ and other consonants

Table 38: Some secondary articulations.

Phonetic term	Brief description	Example symbols
Labialization	added lip-rounding or protrusion	s^w, b^w, t^w
Palatalization	raising of the front of the tongue	s^j, b^j, t^j
Velarization	raising of the back of the tongue	$s^ɯ$, $b^ɯ$, $t^ɯ$
Pharyngalization	retracting of the root of the tongue	$s^ɒ$, $b^ɒ$, $t^ɒ$

Table 39: Contrasts involving labialization in Twi.

càcà	straw mattress	ɔ́c^{w}á	he cuts
ɔ́ɟá	he leads	ɔ́ɟ^{w}á	he carves
ɔ́ɲá	he finds	ɲ^{w}á	snail
ɔ́ʃɛ́	he wears	ɔ́ʃ^{w}ɛ́	he looks at

Table 40: Contrasts involving palatalization in Russian.

brat	brother	brat^j	to take
krɔf	roof	krɔf^j	blood
stal	he has become	stal^j	steel
ʒar	beat	ʒar^j	cook

Table 41: Contrasts involving velarization and/or pharyngalization in Tamazight.

ibð̆at	he began it	ibð̆^{ɒ}at	he divided it
taʒr	rich man	t^{ɒ}aʒin	stew
zurn	they are fat	z^{ɒ}urn	they made a pilgrimage
irsa	he dismounted	irs^{ɒ}a	he quieted down

in many forms of English. For typographic reasons labialized sounds are here specified by a small ^{w} ; it is difficult to put a diacritic under or over a symbol such as ʃ . But it should be noted that not only is there no intention of indicating a sequence, but also the symbol ^{w} indicates simply a labial articulation, and not a labial velar articulation.

We may note here that it may be necessary to consider two different kinds of lip rounding, although the distinction is not one which is strictly relevant to the secondary articulation of labialization. It is possible to form a small lip aperture by bringing the corners of the mouth forward and protruding the lips; or it can be done by closing the jaw and bringing the lips together vertically, so that the side portions are in contact, but there is a gap in the center. Sweet (1890) called these two possibilities inner rounding and outer rounding; Heffner (1950) uses the terms horizontal lip rounding and vertical lip rounding; perhaps a better pair of terms might be lip rounding (which would include protrusion) as opposed to lip compression. Recently Kelly (1966) has shown that the distinction occurs in Urhobo, which has (in addition to labiodental and labial velar fricatives v and w^) both labial velar and labial approximants. Kelly points out that this latter sound, which was symbolized ʋ in the Urhobo examples in Ladefoged (1964), is accompanied by the tongue position of the adjacent vowels; so when it occurs with high back vowels, it is distinguishable from the labial velar approximant w only by the fact that the lip gesture involves (vertical) compression as opposed to (horizontal) rounding.

Several sounds which are often said to be labialized are, from our point of view, sequences of partially overlapping articulations. Thus the beginning of the word *quick* clearly involves a sequence in that the lip rounding in a phrase such as *pretty quick* is not coterminous with the formation and release of the closure; and although both the formation and release of the stop consonant in *see two* may be accompanied by lip rounding, the peak in the labial activity is always much nearer the latter. Both the stop in *quick* and that in *two* are best described in terms of the overlap between successive items, rather than in terms of a simultaneous secondary articulation of the kind that occurs in the initial consonants in *rick* and *ship*.

It may turn out that labialization is the only secondary articulation needed as a linguistic category. The correlates of all the other features listed in Table 38 may be more easily specified in terms of overlapping vowel-like articulations. Thus palatalization can usually be associated with a high front vowel-like articulation which occurs very slightly after the consonant, and which has much shorter durational characteristics than those associated with a normal vowel. Palatalization is a well known feature of Russian and other Slavic languages; examples are given in Table 40.

Velarization and pharyngalization are also vowel-like gestures of part of the tongue, the former being associated with the raising of the back of the tongue, and the latter with its retraction. No language uses a contrast between these two possibilities. Some forms of Polish use a contrast between a palatalized and a velarized l , which may necessitate our regarding these two categories as definitely additive components, not explicable in terms of sequences. But in most other languages a sequential explanation is possible. In Berber languages, for example, the distinction between emphatic and non-emphatic consonants is largely that the former are velarized or pharyngalized, whereas the latter are not; Tamazight examples (suggested by Johnson, 1966) are given in Table 41. But in these languages (and in Arabic) the effect is certainly as noticeable on the subsequent vowel as on the consonant itself, and the secondary articulation cannot be said to be coterminous with the consonant, or with any other segment.

Abercrombie (1967) has suggested that retroflexion should also be regarded as a possible secondary articulation, which might be applicable to vowels only. In some forms of American English, for example, there are vowels which involve the simultaneous curling up of the tip of the tongue, so that there is a retroflex articulation. Phonetically, these sounds can be regarded as retroflex approximants with the addition of the appropriate vowel articulation. If we consider that palatalization, velarization, and pharyngalization may be described in terms of the categories which we will have to set up for vowels, then we need no further categories to specify the so-called r-colored vowels, which, from our point of view, will be vowel-colored r's.

The categories for specifying consonant segments which have been suggested (but not formalized) above, are certainly insufficient; there are a number of languages with consonant segments which cannot be described in terms of only these categories. In Kamba, for instance, the voiced palatal approximant (orthographic *y*) is regularly accompanied by a dental or interdental approximant which is something like a frictionless ð̆ . In a strict sense neither of these articulations is secondary to the other; accordingly it might be considered necessary to postulate an additional double articulation to be called dental-palatal which would be parallel with labial-velar, labial-palatal, and labial-alveolar. Other sounds found in Kom and Kutep may necessitate an additional secondary articulation, elsewhere termed labiodentalization (Ladefoged 1964). And earlier in this book we noted the rare possibility of a linguolabial stop (Lounsbury, personal communication). Many other gestures also occur in ideophones; thus Shona has a voiceless bilabial approximant as in blowing out a candle in a word meaning *it is all over* and a word bam meaning the crack of a rifle, where the vowel is roughened by epiglottal friction (cf. Doke 1931). Phenomena such as these may require some changes or additions to our present categories; but until we have more data it seems better to regard them, like grammatical anomalies, as extraneous events which have to be listed.

Table 42: Some of the vowel contrasts in Ngwe

1.	mbi	white chalk	5.	mbɑ	person
2.	mbe	knife	6.	mbɔ	god
3.	mbɛ	sheath	7.	mbo	hands
4.	mbæ	pepper	8.	mbu	corners

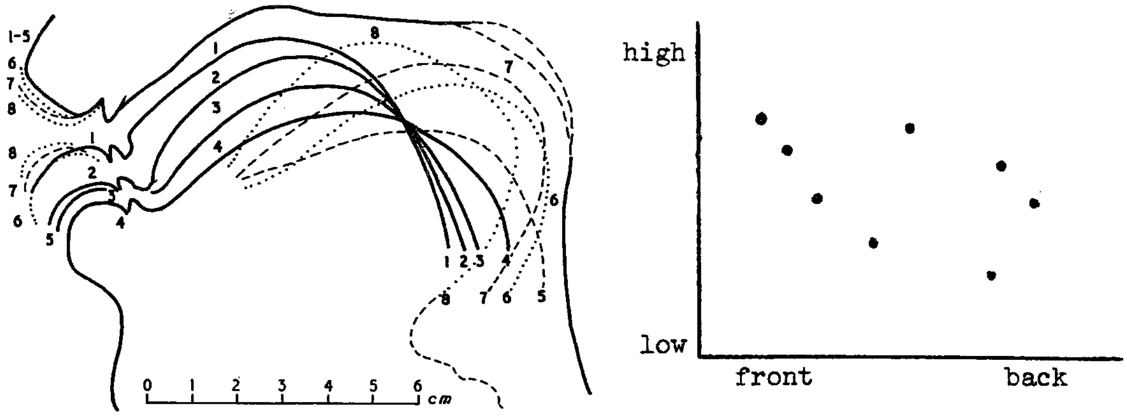


Figure 4: The articulation involved in the eight Ngwe vowels in Table 42. The diagram on the left (from Ladefoged 1964) is based on tracings from single frames of a cine-radiology film. That on the right shows the relations between the "highest points of the tongue" in these articulations.

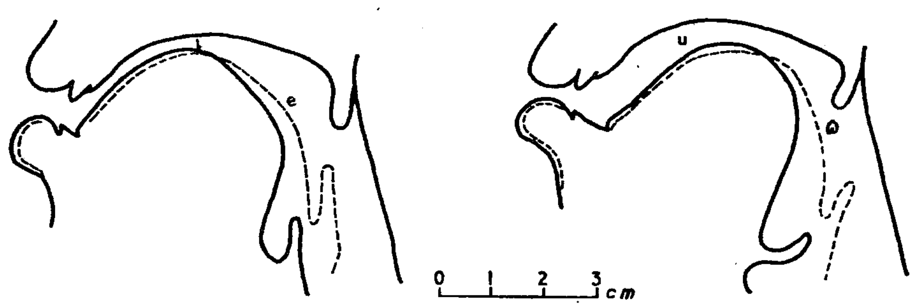


Figure 5: Tracings from single frames in a cine-radiology film showing the articulations involved in the last vowel in each of the Igbo words: óbì , ɔ̀bé , íbu , ɔbɔ̀ heart, poverty, effort, boast. (From Ladefoged 1964)

Vowels

The description of vowels in terms of a limited number of categories raises a number of problems, some of which I have discussed at length elsewhere (Ladefoged 1962, 1967). In general we can say that vowels can be described as points on a continuum in a way that is not true for consonants (with the possible exception of the categories for place of articulation). Our first task is to attempt to define the parameters which specify the vowel continuum. For the last hundred years the traditional way of doing this has nominally been in terms of the position of the highest point of the tongue and the position of the lips. The disadvantage is that the terms are often not in accord with the physiological facts.

We may begin discussing possible articulatory descriptions of vowels by reference to some of the vowels of Ngwe (Dunstan 1964), examples of which are given in Table 42. This is convenient partly because these vowels represent a wide range of phonetic qualities, eight of which are fairly similar to the well known cardinal vowels (Jones 1956), and partly because we have good cineradiology data available (from Ladefoged 1964) which enables us to draw accurate diagrams of the articulatory positions as shown in Figure 4. Phoneticians have not stated how they would locate the highest point of the tongue; the most suitable way would seem to be by presuming that the lower surfaces of the upper teeth form a horizontal plane. (Approximations usually have to be made in that the teeth are seldom on a single plane; but this method works better than using the most nearly flat part of the palate as a reference line, since the majority of subjects have a more curved palate than is the case for this particular subject.) The lower part of Figure 4 shows the relation between the highest points of the tongue determined in this way. It is apparent that the four front vowels lie on a straight line and may be appropriately specified by the traditional labels. But the four back vowels have very different tongue shapes from the front vowels, and these tongue shapes can be considered as differing simply in terms of the single parameter called tongue height only by neglecting large and varied differences in the front back dimension. Moreover, although these four back vowels may not be absolutely identical with the cardinal vowels ɑ, ɔ, o, u they are certainly fairly similar and form a series of approximately equal auditory steps; but the highest points of the tongue are far from equidistant.

The only way of regarding the articulatory positions of the back vowel as being approximately equidistant is by reference to the position of the point of maximum constriction of the vocal tract. This point, which was suggested as a reference point by Stevens and House (1955), gets progressively further away from the glottis by roughly equal steps on a logarithmic scale as one goes from ɑ to u . It is thus an appropriate

Table 43: Contrasts involving front rounded vowels in French.

vi	life	vy	seen	vu	you
de	thimble	dø	two	do	back

Table 44: Contrasts involving back rounded and unrounded vowels in Mandarin Chinese.

tsʰɯ	times	sɯ	four
tsʰù	vinegar	sù	speed

Table 45: Examples of contrasts among high and central vowels in Ngwe. I am not sure if the word meaning brass gong should contain or (a high or a mid back unrounded vowel); it is, however, quite clearly not a central vowel.

aty	stick	mbi	white chalk	mbɨ	dog	mbu	corners		
								mbɤ	brass gong
				ntsʌ	water				

Table 46: Contrasts involving high rounded vowels in Swedish.

vy:	view	hʉ:s	house
sy:n	sight	nʉ:	now
sø:t	sweet		
rø:d	red		

way of specifying back vowels, just as tongue height provides a useful description of front vowels. But specifications of tongue shape in terms of the position of the point of maximum constriction may sometimes be misleading. There is, for instance, no articulatory or acoustic discontinuity corresponding to the discontinuity in this form of specification which occurs when one goes from ɛ (in which the maximum constriction is near the hard palate) to æ (in which it is nearer the pharynx, see Figure 4). There seems to be no single simple set of parameters which is equally appropriate for specifying the tongue shapes of all these vowels.

The vowels shown in Figure 4 are in no way exceptional; and there are similar difficulties in specifying the position of the highest point of the tongue in the only published X-ray data on a complete set of cardinal vowels (Jones 1929). In these vowels, as in the Ngwe examples, there are both great differences in the general shape of the tongue in the set of front vowels and in the set of back vowels, as well as anomalous positions of the highest point of the tongue in back vowels. 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.

Some phoneticians have suggested additional articulatory parameters for specifying the shape of the tongue, such as narrow and wide (Sweet 1890) and tense and lax (Jakobson and Halle 1964). There are certainly grounds for believing that the highest point of the tongue or the point of maximum constriction of the vocal tract (it makes no difference which we consider in this discussion) can be in a given place, but the tongue can be more or less bunched up lengthways (in the anterior posterior dimension). Sweet pointed out very clearly that it is possible to produce a particular height of the tongue with either a certain jaw position and a bunched up tongue, or a higher jaw position and a relatively flattened tongue. Figure 5 (adapted from Ladefoged 1964) shows the tongue positions in two pairs of Igbo vowels which are distinguished in a similar way by the extent to which the root of the tongue is pulled forward.

Finally in considering the articulatory parameters of vowels we must note that the degree of lip rounding is an independent variable. There is a tendency in the languages of the world for front vowels to be unrounded and back vowels to have lip rounding increasing with tongue height as in the primary cardinal vowels. But a great many languages have front rounded vowels (usually with greater lip rounding for higher vowels); and several have back unrounded vowels (usually with greater lip spreading for the higher vowels). Examples of front rounded vowels in French are given in Table 43; and back unrounded vowels in Mandarin Chinese are exemplified in Table 44. Malmberg (1956) has suggested that the difference between what we have called (horizontal) lip rounding with protrusion and (vertical) lip compression may be contrastive in Swedish.

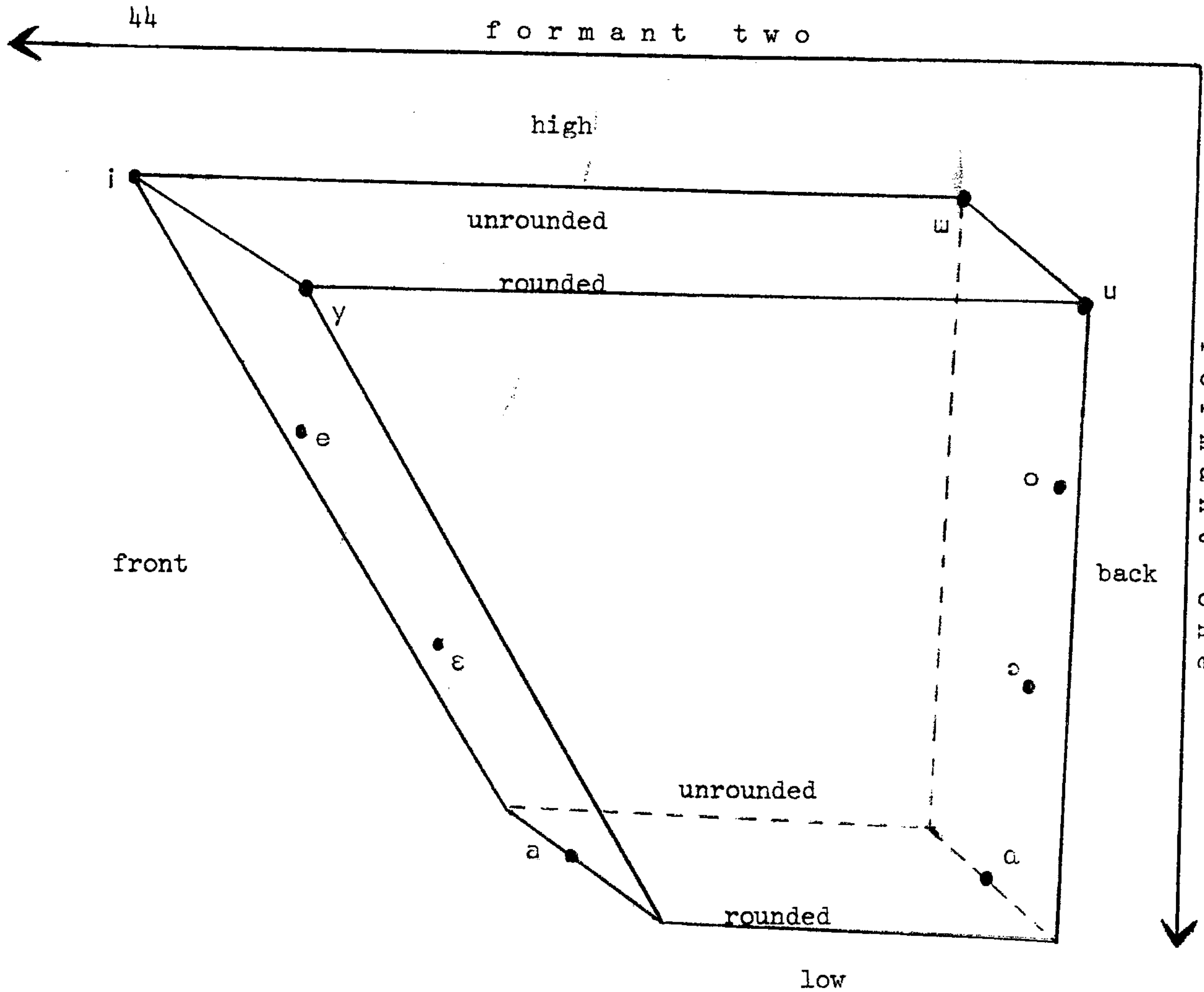


Figure 6: The (pseudo) articulatory vowel continuum, and its relation to formant frequencies.

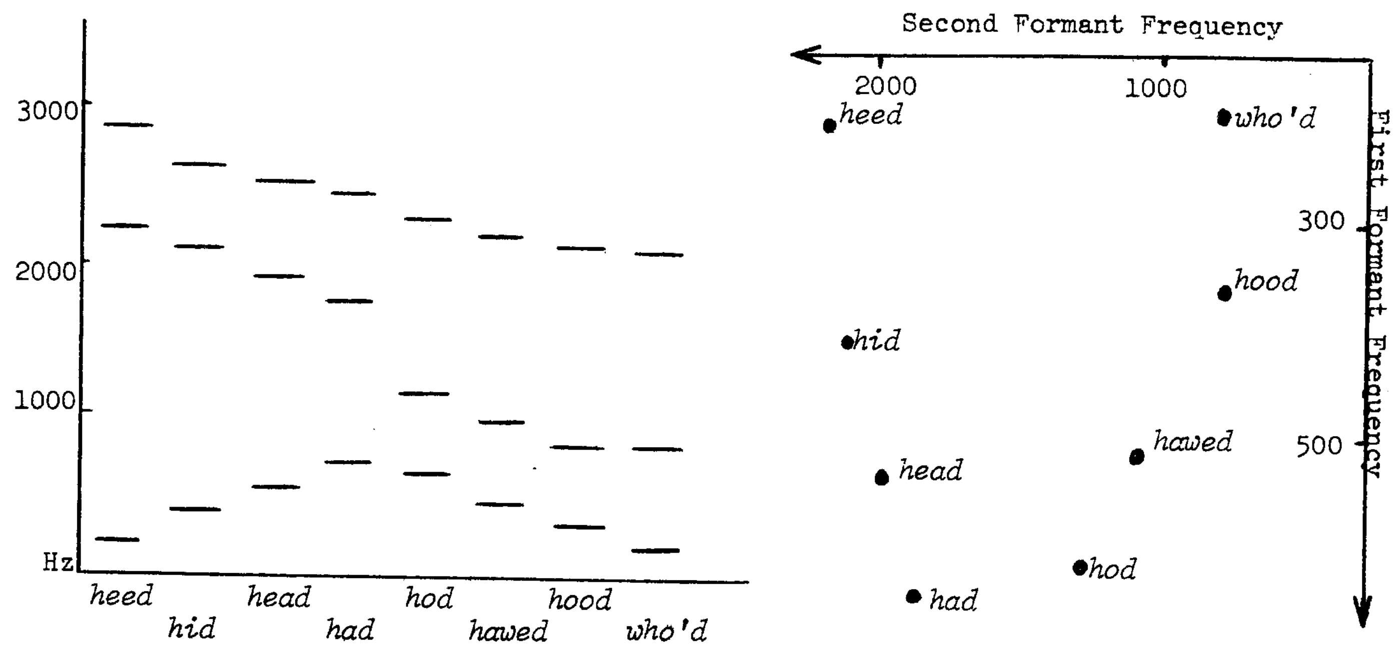


Figure 7: A schematic spectrogram showing the frequencies of the first three formants of the vowels in some English (RP) words; and a formant chart showing the relation between the first and second formants in these words.

A convenient way of diagramming the vowel continuum is shown in Figure 6. It may be seen that the cardinal vowels do not lie on a single surface of this space; and we must also remember that the terms for the tongue positions are simply traditional labels, which are not easily correlated with the articulatory facts. The three dimensional continuum has been drawn as if viewed from this particular angle because the individual vowels may then be more easily correlated with the acoustic parameters which are indicated on the right of the figure. So far in this survey of sound types we have not been concerned with acoustic descriptions. Later we will have to consider whether acoustic descriptions of consonants lead to a more explanatory theory of phonetics; but a more traditional physiological approach was sufficient for a first account of the contrasts which a theory of phonetics must encompass. In descriptions of vowels, although a pseudo-articulatory terminology may provide an adequate set of labels for auditory descriptions, we have seen that we do not have, as yet, a set of articulatory parameters which will specify vowel quality. Accordingly we must consider whether the vowel continuum can be better described in terms of acoustic parameters.

The basic acoustic data are the frequencies of the formants which characterize each vowel. Roughly speaking, we can say that the sound of a vowel consists of the pitch on which it is said (which is due to the rate of vibration of the vocal cords) and the pitches of the two or three principal groups of overtones (which can be associated with the resonant frequencies of the vocal tract). These groups of overtones are called formants; they are the principal determiners of vowel quality. (See Ladefoged (1962a) for an elementary account of formants and acoustic phonetics.) The frequencies (roughly pitches) of the formants of the vowels in some English (RP) words are shown schematically on the left of Figure 7. If these vowels are whispered (so that there is no pitch which can be associated with the action of the vocal cords) the falling pitch of formant two can be heard quite easily. If they are said with a creaky voice quality (laryngealized phonation) the rising then falling pitch of formant one is often apparent. The variations in formant three can not be demonstrated in any simple way.

A common method of plotting formant frequencies is shown on the right of Figure 7. Sometimes only the first two formants (F1 and F2) are represented as here, and sometimes more elaborate combinations involving the third formant (F3) are plotted, such as (F1 + F3) against (F3 - F2). But apart from the work of Jakobson, Fant and Halle (1951), which will be discussed shortly, there has been very little attempt to locate three or more independently varying acoustic qualities of vowels; and all the two dimensional plots lead to difficulties. There is no way of giving a two dimensional specification of front rounded and back unrounded vowels which does not involve some confusion with centralized vowels with other lip positions. If the first formant is plotted against the second the point for the secondary cardinal vowel y will be very close to that for the vowel in the English word *hid*, as might be

Figure 8: The dimensions of vowel quality. Where there are two symbols in an undivided box the one nearer the center of the chart represents a lax variant.

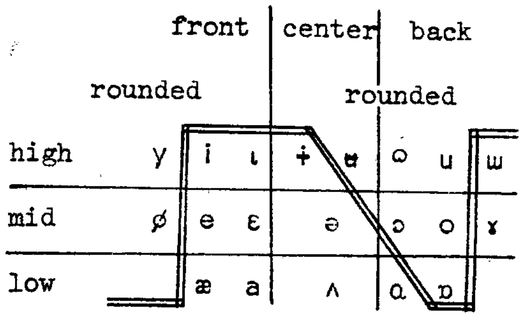


Table 47: Further examples of the oppositional use of some of the symbols in Figure 8. The words above the line are in Highland Scottish English; those below are in English RP.

bit (beet)	bɩt (bit)		kɷd (could)	but (boot)
bet (bait)	bɛt (bet)	kət (curt)	kɔt (caught)	bot (boat)
kæt (cat)		kʌt (cut)	kɑt (cart)	kɒt (cot)

inferred from the view of the vowel continuum shown in Figure 6.

It is probably impossible to classify all the vocalic oppositions within the languages of the world in terms of only two parameters. But most linguists and phoneticians seem willing to describe the majority of vowels in terms of three or four dimensions, which they designate by articulatory labels, but which it would be better to consider as auditory qualities (cf. Ladefoged 1962, 1967). The most usual terms for these dimensions are: lip rounding; front - back; high - low (or close - open); and (recognized by some investigators only) tense - lax. For the moment we will assume that these auditory parameters are those we require for categorizing phonemic contrasts among vowels. Our next problem is to consider how many possible contrasts there are within each dimension. A tentative solution is indicated by the array of symbols in Figure 8. The claim being made is that if we recognize the binary dimension of vowel tension, operating only for front rounded and back unrounded vowels, then we need to recognize only three degrees of vowel height and three degrees of vowel fronting; in the cases of front rounded and back unrounded vowels, only two degrees of vowel height need be recognized; and in the cases of front and back vowels and high central vowels, two degrees of lip rounding must be recognized.

It should be remembered that the terms vowel height, fronting, tension, rounding, etc. are being used as labels for features of auditory dimensions for which the acoustic and physiological correlates have not been given. This is an unsatisfactory, interim solution, which we hope will be remedied soon. But its advantage is that it works; linguists seem to be able to distinguish three degrees of vowel height and fronting, and two degrees of tension and rounding; and there are seldom disagreements over which symbol (out of this limited set) should be used when specifying the relative qualities of an opposition. (Of course if we wish to do more than specify the relative qualities in a phonemic contrast, then we will have to use some other system, such as the cardinal vowel system (Jones 1956) which aims at specifying absolute phonetic quality but this is irrelevant to our present purposes.)

The matrix in Figure 8 is similar to one suggested by Stockwell (1959) for use in transcriptions of dialects of American English. The major differences are the addition of front rounded and back unrounded vowels (which were not needed for Stockwell's purposes), and a slightly different placing of the symbols, so that the symbols are arranged more in accordance with the *Principles* of the International Phonetic Association (1949) as opposed to the practice of American dialectologists. The matrix in Figure 8 has only three degrees of vowel height, and it is therefore impossible to make it coincide exactly with IPA usage which has four. The arrangement shown seems the best possible compromise. Stockwell suggests that his matrix allows for the transcription of all the differences that have been discussed as possibly phonemic in some dialects of American English; and he gives examples which illustrate the necessity of a matrix of this size, with six front vowels and six back vowels. I would imagine that the categories shown are probably sufficient to account for the vocalic oppositions which occur within each of the languages of the world. In fact, I have not myself heard contrasts between each of the adjacent items. Thus, although I have heard varieties of English which have a tense æ which is different from the lax a found in other dialects, nevertheless I do not know of a language which uses this difference in a phonemic contrast. Nor do I know of any use of the three central unrounded vowels ɨ , ə , ʌ within one language; and I am not sure if there are examples of contrasts between the three unrounded mid vowels e , ə , ɤ . Ngwe, however, has contrasts between a number of high and central vowels as shown in Table 45; and Swedish has some rounded central vowels as shown in Table 46. Further examples of the oppositional use of the symbols in Figure 8 are given in Table 47.*

Syllables

So far we have discussed categories for specifying vowels; but we have not considered criteria for distinguishing between vowels and the class of consonants we called approximants. There are, in particular, the so-called semivowels j , w , ɥ which are very similar to the corresponding vowels i , u , y . (See Tables 25 and 43 for examples of all these sounds in French.) It does not seem possible to distinguish between these two groups of sounds by adding an extra category based on articulatory criteria, such as the degree of articulatory stricture, or the rate of change of articulatory position. It is not always true that vowels have a more open articulation than semivowels; and on some occasions semivowels have a steady state portion which is comparatively long (Lehiste 1964), and may be longer than that in some vowels. A better solution (basically that of Pike, 1943) is to distinguish between vowels and semivowels by calling the one group syllabic approximants, and the other non-syllabic approximants. We shall in any case need to distinguish between syllabic and non-syllabic sounds in other cases, such as the l sounds in *coddleing* and *codling*; in my speech these words differ by having an identical number of segments (there is no extra vowel in the first word) with identical properties, except for the syllabicity of the l . (It is, of course, irrelevant that at some level of phonological abstraction there might be an extra vowel represented in the first word; the final phonetic representation must be able to characterize the difference in the way described above.)

The difficulty with positing the categories syllabic and non-syllabic is that there are no simple physiological or acoustic properties clearly distinguishing these two possibilities. But although there is no single muscular gesture marking each syllable (Ladefoged 1958, 1967), we still need a physiological unit of this size to account for the timing and coordination of the articulatory movements. There is evidence

---

*Note that the Ngwe vowels given in Table 42 and Figure 4 are specified in terms of the traditional cardinal vowel system, with æ being used in place of a , since the Ngwe vowel is not as low as possible. It may be a weakness of Figure 8 that it does not allow for four degrees of so-called vowel height, as may be required in Ngwe.

Table 48: Contrasts involving long and short consonants in Italian.

fato	fate	fat:o	done	kade	he falls	kad:e	he fell
fola	fable	fol:a	crowd	nono	ninth	non:o	grandfather

Table 49: Contrasts in vowel length in Kamba. The four degrees of length are represented as: a a• a: a::

1.	kwele la	measuring	koʃa	start
2.	kwele•la	moving backwards and forwards		
3.	kœle:la	aiming at	koʃa:	giving birth
			koʃa::	giving birth frequently

(Koshevnikov et al., 1965; Ladefoged 1967) that speakers organise the sequences of complex muscular events that make up utterances in terms of a hierarchy of units, one of which is of the size of a syllable; and it is certainly true that speakers know how many syllables there are in an utterance.

Length

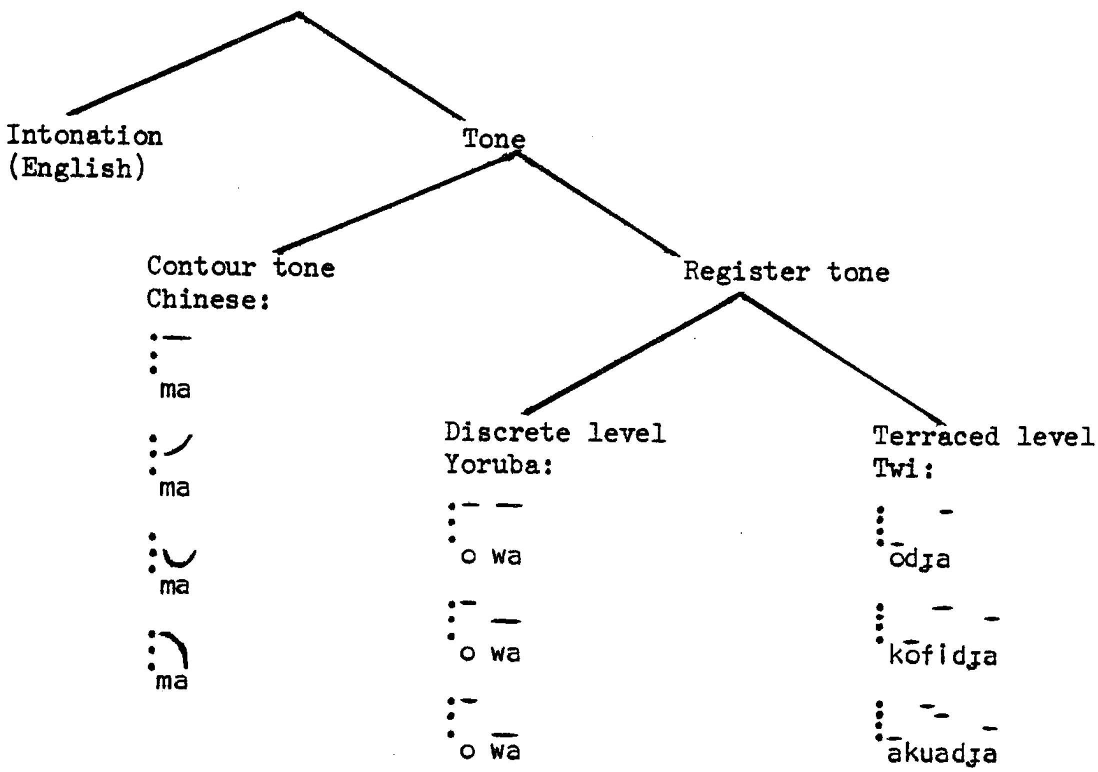
Many languages use contrasts in the lengths of segments. The abstraction is straightforward and easily specifiable in measureable physical terms. It is, of course, a relative quality (as are many other phonetic qualities) much influenced by rate of utterance (which itself may be partially determined by the age and emotional state of the speaker). Comparatively few languages use contrasts between long and short consonants within the same morpheme; Italian is a well known example (see Table 48). But many languages have contrasting long and short vowels. In some Bantu languages vowel length may be used to mark not only lexical oppositions but also other grammatical contrasts. This possibility leads to there being possibly four degrees of vowel length in Kamba, three of which are clearly contrastive as shown in Table 49 (elicited from examples in Whitely and Muli, 1962).

Pitch

If pitch is specified in terms of the actions of the laryngeal muscles which control the rate of vibration of the vocal cords (or the corresponding phenomena on a physiological type speech synthesizer), then we may begin by considering it independently of any of the categories mentioned so far. An item indicating a particular contrastive pitch can be considered to apply until the occurrence of the next pitch indication, irrespective of the articulatory items which may be occuring.

All languages use variations in pitch to convey differences in meaning. Table 50 summarizes and illustrates the principal methods. The first division is into the use of pitch for conveying syntactical information (commonly called intonation) as opposed to lexical information (commonly called tone). Thus English varies the intonation of a sentence or a clause to produce differences such as those between a statement or a question, and Chinese varies the tone to produce different lexical meanings. Pike (1945) has suggested a division into contour tone languages such as Chinese, in which the various tones include some in which the essential feature is a changing pitch, as opposed to register tone languages, such as most African languages, where the tones marking the lexical items are comparatively steady state. Welmers (1959) has suggested a further division of register tone languages into discrete level languages, such as Yoruba, which have tones in which the pitches are always in the same relation to one another (e.g. one high, one in the mid range, and one low), as opposed to terraced level languages, such as Twi, in which the pitch of one of the tones depends

Table 50: The linguistic use of pitch.



Combinations of Tonal and Intonational Features:

1. Tone on all items; intonation in parts of some utterances.

   e.g. 'final lowering' in Yoruba: in environment # high / mid / low →

2. Tone on all items; intonation throughout.

   e.g. 'downdrift' in Hausa: malam insu yana ba su nama
   'teacher their he gives them meat'

   note also question: malam insu yana ba su nama

3. Tone on some items; intonation throughout.
   e.g. Swedish, Gaelic (Lewis);

   Fula: o waddii ceede den hande
   'he's brought money the today'
   o waddii ceede fun hande
   'he's brought money all today'

4. No tone; intonation throughout.
   e.g. English

on which of the other tones occurs immediately before it. Thus in the Twi example in Table 50 (based on Schachter 1961) the stem of the word meaning *fire* has the highest possible pitch when it occurs after a low pitch (as it does when it has only a nominal prefix), but when it occurs after a high pitch (as in the second syllable of the name *Kofi*) it is one step lower; and when it occurs after a tone which is itself a step down from the highest possible pitch (as in the third syllable of the name *Akua*) then it is one step lower still.

In practice languages always have a mixture of these methods of using pitch (cf. Gleason 1961a). There is probably no language, however tonal, which does not have some intonation features corresponding to a grammatical unit such as a clause or a sentence. As shown in the lower part of Table 48 tone languages such as Yoruba have merely some form of lowering of the final tone; but other tone languages such as Hausa, have a falling or a rising intonation pattern over the whole sentence, such as in English. My own experience with tone languages has been restricted to those contrasting at the most four different tones. But Trique is said to have five phonemic levels of pitch (Longacre 1952); and there are reports of languages with six tone levels within a single tone bearing unit.

Opinions differ on the specification of intonation languages. The major apparent disagreement is over whether intonation should be specified in terms of a number of pitch levels (Pike 1946, Trager and Smith 1951), or in terms of a number of contours or tunes or configurations (Jones 1956, Halliday 1962, Bolinger 1958). In fact it seems clear that from the point of view of the higher level phonological rules, the complete contours contrast with one another; but the phonetic specification must be in terms of target pitches. In other words, it may be said that in English, for instance, we have five significant intonation contours (Halliday 1962) so that we can contrast a single sentence such as *Yes* with a falling intonation (meaning *I agree*), with a similar sentence with a low rising intonation (meaning *I am listening, carry on*), with another with a high rising intonation (meaning *Did you say 'yes'?*), with one with a rising falling rising intonation (meaning *I am doubtful*), and with one with a rising falling intonation (meaning *I am certain*). But when it comes to specifying the action of a speaker, these five English intonation contours must be reinterpreted in terms of target pitches. The relation between intonation contours and target pitch levels is in some ways (but not in all ways) analogous to that between phonemes and the bundles of distinctive features or simultaneous categories of which they are composed. The contours which occur in languages with lexical tones must also be specifiable in terms of target pitch levels. Chao (1930) finds that he needs five levels to specify the starting, ending, and turning points in the four tones of Mandarin Chinese. At the moment I can do little more than guess at the total number of pitch levels which may be needed for specifying the languages of the world. But experience with operating speech synthesizers by rule would seem to indicate that we do not need to have more than six target pitches in order to specify all the pitch contrasts which occur in any language. We may be able to do with as few as four.

Stress

Stress is more difficult to handle, not because the phonetic correlates of stress itself are more complex (as we have seen, in terms of the physiological specification, it involves increased activity by the respiratory system producing a greater subglottal pressure), but because many allophones of the articulatory segments are organized in terms of their position within the stress group--the phonological unit defined by the onset of stress. For example, the difference between the pair of phrases *a stray tissue* and *a straight issue* is not only in the timing of the increase in subglottal pressure, but also in allophonic differences such as the use of an aspirated t in the last word of the first phrase. We are not, of course, concerned with whether the position of the stresses can be predicted from a knowledge of the syntax and morphology as suggested by Chomsky (1964), just as it is irrelevant whether intonation contours are predictable from grammatical rules distinguishing between different types of clauses. Both stress and pitch are clearly the phonetic correlates of certain linguistic contrasts, which may, at a higher level, be specified simply in terms of the surface structure of the sentence (Chomsky and Halle, forthcoming). But ultimately it is irrelevant to the phonetic specification that these contrasts may be syntactic or morphological rather than lexical.

The number of stresses which we need to recognize poses problems similar to those in deciding on the number of target pitch levels. My own tentative solution is based solely on the analysis of English--the only language for which I have sufficient data concerning the peaks in subglottal pressure. My impression (which is far from fully substantiated) is that the degree of stress can be specified as a variation in the subglottal pressure by stating only two target levels: there is a target associated with each stress group (roughly the occurrence of each primary or secondary stress in the Trager and Smith (1951) system), and an additional target associated with the peak of each intonation contour. We can call these the lexical and the syntactical targets. Also associated with each intonation contour is a general downdrift so that stresses of any kind at the beginning of a sentence are larger than those nearer the end. Schematic curves illustrating the additive principle are shown in Figure 9. Actual subglottal pressure records which may be interpreted in this way were first given in Ladefoged (1963) and repeated in Ladefoged (1967).*

*New data indicate that there may have to be some revisions in this paragraph. Figure 9 is omitted for the moment.

2

PRELIMINARIES TO A THEORY OF PHONETICS

Phonemics and Systematic Phonetics

The sounds of one language are often difficult to describe in terms of the categories used for describing another; and the more one looks at the languages of the world, the more one seems to have to increase the number of phonetic categories required for making adequate descriptions. Whether this is so or not depends in part on what one means by making an adequate phonetic description. Now we will extend the discussion of the previous section and assume that an adequate phonetic description must be able to specify both the linguistic oppositions (such as the phonemes) which occur within a language, and also the characteristics of that language as opposed to other languages. Given this we may then ask whether we can set a limit on the number of phonetic categories required for describing the languages of the world. It seems that we can do so provided that we use the categories to specify only the linguistically relevant sounds within a language. Thus, to take an example, the English b in *bat* is different from the French b in *bête*, in that in French the vibrations of the vocal cords continue throughout the stop closure, whereas in English they often do not. But we do not need to specify these two b sounds in terms of different categories as long as we are limiting ourselves to specifying only the oppositions within each language.

This is the essential insight behind Jakobson's concept of distinctive features. In a long series of publications (Jakobson 1962, Jakobson, Fant and Halle 1951, Jakobson and Halle 1956) he and his colleagues developed definitions of about 14 binary oppositions which they thought sufficient to characterize all possible phonemic contrasts. Although their categories may not be altogether appropriate, the concept of a distinctive feature specification has been shown to be remarkably useful (e.g. by Halle, 1964, and by Chomsky and Halle, forthcoming). Provided that we can find a rigorous way of deciding when an opposition in one language (such as p-b in English) can be equated with a similar opposition in another language (such as p-b in French) then a definition of a limited set of features seems not only practical but also desirable as part of a general theory of language.

Nearly all theories of linguistic description have assumed that it is possible to describe languages in terms of a set of general phonetic categories. But we should note that in the works of the older great phoneticians such as Bell (1867) and in the pre-phonemic era of the development of the IPA, there was often no clear distinction between categories necessary for specifying the oppositions within a language, and the additional theoretical apparatus required for characterizing that language as opposed to all others. Even at a much later date, Joos (1957) and other linguists seem to maintain that there should be no such distinction and to advocate *ad hoc* description for each language; Hockett (1955) has said that "it is impossible to supply any general classificatory frame of reference from which terms can be drawn in a completely consistent way for the discussion of every individual language." But, as has been frequently pointed out (Jakobson and Halle, 1956; Chomsky and Miller, 1963), all linguists (including those cited above) need phonetic descriptions which imply some absolute frame of reference. Unless we are able to use categories such as alveolar and bilabial, which refer to observable phenomena, we cannot describe the first and last sounds in *bib* as being in some way the same, and also know that the first sound in *bib* is not to be identified with the last sound in *did* (cf. Fischer-Jørgensen 1952). Furthermore, when we compare languages at the phonological level we must have some common frame of reference. Otherwise we could do no more than count oppositions, without even knowing what proportion of the oppositions were between vowels and what between consonants in each language. Of course, it is possible to take the view that each language has to be described as an entity in itself, and that we can never compare languages. I would rather hold the view that it is possible to compare any two items in the universe--if we take a sufficiently large framework. It may be true that the framework required for comparing languages has to be very large and very complicated; but our job as linguists requires us to define it. The position taken here is that the frame of reference is much simpler when we separate out the task of specifying the oppositions within a language from that of characterizing phonetic differences between languages.

An explicit way of making this distinction has been proposed by Halle (1964a) He suggests that it is possible to use the Jakobsonian distinctive features (which he assisted in developing) as a set of binary categories for classifying the oppositions within a language; and then, at a later stage in a series of rules, give a more detailed phonetic account of the characteristics of the language under description by replacing the binary categories with a quantitative specification of the relative values for each feature. In the case of the differences between b in English and French, we would therefore begin by using categories which distinguish the oppositions between each of these sounds and all other contrasting items within each language, before going on to consider what it is that characterizes English b as opposed to French b . The first step might involve stating whether each sound was, or was not, more voiced, bilabial (or grave), and interrupted (like a stop) than the contrasting sounds; and the second might involve stating the degree of voicing (and other concomitant features) present in each case.

It is often useful to classify the sounds of a language in terms of such general categories at an early stage in the description; the classificatory categories may then be used for dividing the linguistic oppositions within each language into a number of natural classes, the members of which operate in the same way from the point of view of the grammar of the language. Thus, in English, we need a class of voiced sounds as opposed to voiceless sounds so that we can make simple statements about the form of the plural suffix in nouns. But if our description of English is to be complete, we must, at a later stage, specify how much voicing there is in each sound in each circumstance, and thus characterize precisely what it is that makes an Englishman sound like an Englishman even when we are paying no attention to the meaning of what he is saying. As Firth (1957) has said "it is part of the meaning of an American to sound like one." This is an unconventional use of the word meaning, which has not always been understood by Firth's detractors (Langendoen 1964). It would probably be more acceptable if we rephrased the same thought in current terms, and said that a linguistic theory should be able to characterize both the oppositions within a language (the differences between the members of the set of all possible sentences) and the contrasts between languages (all and only the features which mark the sounds of the language as being different from the sounds of other languages). One of the objects of this book is to assess the extent to which it is possible to construct a theory of phonetics which will achieve this goal.

Others have suggested different goals for a theory of phonetics. Thus, for the traditional American linguists, phonetic analysis of a language had to precede its description in linguistic terms; and it had to be made completely independently of any knowledge of the oppositions which occurred. A satisfactory way of making phonetic descriptions for this purpose may be theoretically conceivable, but it would be very complicated. The best attempt to date is probably that of Peterson and Shoup (1966). However, Chomsky (1964) has shown that it is not possible to describe a language adequately by starting from descriptions of sounds without reference to their linguistic function. So we do not need a theory of phonetics which will include a procedure for classifying sounds irrespective of their function within a language. Much to the discomfort of some phoneticians (and some linguists), phonetics is not a science that linguistics must presuppose. All that is necessary for linguistic purposes is a theory of phonetics which will allow us to account first for the oppositions within a language, and then for the relations between languages.

Physical Phonetics

But if it is to be interesting, the description of each language must also be testable; and the possibility of making a sufficient test must be inherent in the underlying theory. The difficulty is that languages are abstractions. What we call a language such as English or French is a code that enables individual Englishmen and Frenchmen to communicate with one another. But we cannot test descriptions of a code without reference to its manifestations; the only data we have for checking our descriptions of a language are the utterances of individual speakers. It seems, then, that there are three stages which a theory of phonetics must be capable of handling. First, it must permit the oppositions within each language to be specified; this is what Chomsky (1964) calls systematic phonemics. Secondly, it must provide a way of accounting for the particular characteristics of each language; this might be systematic phonetics. Thirdly, it must lead to the specification of actual utterances by individual speakers of each language; this is physical phonetics. Linguistic descriptions which do not meet all three of these requirements are apt to be trivial. In practice the first step involves allocating sounds to contrasting categories, the second to designating relative values of each category, and the third to interpreting these values in terms of measureable units. Thus, to continue our example, both English and French may need the categories voiced and voiceless; but English initial voiced stops may be said to have 15 degrees of voicing, whereas French may have 60; and a particular English speaker may make a given word with a voiced initial stop in which the voicing lasts 15 ± 2 msecs, whereas a particular French speaker may, in his circumstances, have an initial stop in which the voicing lasts 60 ± 2 msecs.

We may now consider the general form of the kind of phonetic description that is being proposed here. It must, like other parts of the description of a language, be capable of being expressed completely in a set of explicit statements or rules, so that we can be sure that no intuitive (possibly fallacious) concepts are required for its interpretation. Ultimately it would be convenient if the rules produced a set of signals which could control a speech synthesizer. Then we could be certain that the entire account of a language was contained in the rules and the theory (which would have to include a specification of the speech synthesizer). Such a description could, in a very literal sense, be part of a generative grammar; and the grammar would be very powerful in that it would contain rules which were not merely possible (specifying correct but not necessarily all the phonetic correlates) but necessary and sufficient (containing all and only the information required to generate speech).

Segments and Targets

One of the major difficulties in achieving this kind of description is in relating the essentially continuous nature of speech with the essentially discontinuous nature of a linguistic description. All linguistic descriptions involve segmentation of some kind, since they all distinguish between an infinity of possible sentences by specifying different arrangements of a small number of discrete units. Virtually the same problem arises irrespective of whether we are attempting to account for what happens in the juxtaposition of segments of the size of a phoneme or a syllable or any other linguistic unit. One solution is that we should specify the ideal form for each unit (or, more precisely, for each category within each unit), and then provide a rule or a set of rules which will specify the extent to which this form is missed because of the influence of the adjacent items. Our description of a language would then include rules for generating a table of values specifying these ideal forms, and rules accounting for the partial overlap or way of getting from one sound to another. The table of values might be expressed in terms of numbers representing relative values of parameters for synthesizing speech, such as formant frequencies and durations; or conceivably it might be in terms of Jakobsonian distinctive features, in which case it would occupy a place in the description somewhat analogous to Halle's descriptive phonetic matrix. Halle has not (to my knowledge) considered the form of the rules for turning phonetic segments into continously varying parameters; but a number of other investigators (Lindblom 1963; Ohman 1964; Holmes, Mattingly and Shearme 1965) have suggested parts of possible schemes. The account which follows is derived mainly from the work of Lindblom and Ohman.

The simplest way of understanding conjoining rules is to consider data from cine-X-ray pictures of speech. It appears that in the middle of the consonant closure in a phrase such as *a key* the tongue is in a position as shown on the upper left of Figure 10; whereas the position at the corresponding time in *a car* is as shown on the upper right of the figure. The positions for the middle of the vowels are shown on the next line of the figure. It is obvious that the positions in the consonants are strongly influenced by those in the vowels. In fact the positions in the consonants are probably predictable from knowledge of an ideal form for the consonant k and an ideal form for each of the vowels i and a, together with some kind of weighting function which specifies the degree to which the position of each part of the vocal tract is absolutely determined by the ideal position. This possibility is indicated in the lower part of the figure, where a guess is made at the ideal or target positions for the vowels and for the consonant. The thickness of the solid lines indicates the extent to which that part of the vocal tract is determined by the ideal forms. Thus the velar area is largely determined during the consonant k by the ideal form; but, even at that time, the tip and root of the tongue are more or less free to be determined by the adjacent item (in this case, the vowel). During the vowel most of the tongue is determined to an equal extent by the ideal form, so in practice it is only in the velar region that the position of the tongue deviates from the ideal position at that time.

Ohman has suggested a mathematical formulation of this kind of conjoining rule, which seems a plausible way of specifying the position of the center of the tongue (the view derived from an X-ray) in vowel-consonant-vowel utterances. He treats the midline of the tongue as being specified by a number of points, each point being a certain distance from a fixed point such as the glottis. Figure 11 shows an instant in the middle of each segment in an appropriate utterance, and a simplified version of Ohman's equations for determining the position P_x of a point x mm from the glottis at each moment in time. Ohman assumes that the tongue starts from the ideal position for the first vowel and finishes at that for the second. This is obviously a simplification of what actually happens, since even at these times the position of the tongue is determined by the adjacent items, the pauses which occur before and after the utterance. But the general principle may still be valid. The main hypothesis is that the consonant gesture is superimposed on what would otherwise be the movement from the first vowel to the second. Equation (2) defines this movement by stating that, at the beginning of the utterance, when $q = 0$, $V_x = V1$ (the ideal position of the first vowel); and that the position V_x gradually becomes more like V2 (the ideal position of the second vowel) as q tends to 1. Equation (1) states that at the beginning of the utterance, when $k = 0$, the position of each point of the tongue, P_x, is the ideal form for the vowel, since the added term is zero at that time. As the middle of the consonant approaches and tends to 1, then the right hand term becomes more important; this term arranges for the position P_x to be that of the ideal form for the consonant if w_x, the factor determining the degree of freedom of articulation, is 1. When w_x is less than 1 the position is somewhere between the ideal form for the consonant and that for the vowel.

It should be emphasized that all this is vastly over-simplified: The ideal forms and weighting functions indicated in Figure 10 are just intelligent guesses; But Ohman (1967) has shown that a theory of this kind can account for actual cine-radiographic data; and it seems that in some such way, with the aid of a table of values for the ideal form of each category within each segment, we should be able to go from a description of speech in terms of discrete segments to a specification in terms of continuously varying parameters.

This approach forces us to consider the circumstances in which we will consider a speech sound to consist of a single segment or target, as opposed to those which require us to specify it as a sequence of two targets. It is at this point that we must remember two of the requirements of a theory of phonetics. It must provide a set of categories for use in descriptions of the phonology of a language; and it must also provide for the interpretation of these categories in non-linguistic terms. There is a conflict between these two requirements when we come to consider items such as affricates. On the one hand these items clearly consist of a sequence of two items, stop and fricative, which we must consider as having independent status in some circumstances. But on the other hand, affricates often have to be considered as single phonological units. In languages such as Hindi and other Indo Aryan languages, and Makua and other Bantu languages, there are contrasts between aspirated and unaspirated affricates (cf. Sindhi examples in Table 16). If we regard affricates as sequences we have to have extra statements (phonological rules) which allow for the possibility of fricatives being aspirated only when they are preceded by a stop within the same syllable. Similarly, there are many languages (e.g. Quechua) where there is a series of ejective stops which includes an affricate; and we do not want to have a special rule saying that fricatives can be ejective only when following a stop which is accompanied by an upward movement of the closed glottis. Clearly, we have to provide for the possibility of a fricative release being part of the same segment as the stop.

Further examples of this kind of problem are not hard to find. In English, laterally released stops (as at the ends of the words *riddle* and *little*) are sequences of stop plus lateral; but in Navaho and other languages, where tɬ' forms part of the series of ejective stops, it is more convenient to regard the lateral release as part of the same segment as the stop. In Tiv (cf. Table 11) there is a series of pre-nasalized stops contrasting with other stops and nasals. This phenomenon, which is common in Bantu languages, leads us to consider stops with and without nasal onsets as single units. Similarly the insistence on simultaneity in secondary articulations may seem ill-advised on phonological grounds. Many languages (e.g. Twi) have sets of consonants which differ in that one set has a slightly subsequent but considerably overlapping feature of labialization or palatalization. The rules accounting for the phonological patterns within these languages will be considerably complicated by a specification in terms of sequential items.

Any description of linguistic forms must ultimately be interpreted (by a human or a machine) in terms of a sequence of items referable to a table of values corresponding to the targets and durations of each item, and conjoining rules specifying the extent to which each target is missed because of the influence of the adjacent items; and it will be much easier to operate such a system if we maintain a clear distinction between those items which involve simultaneous targets, and those which involve sequential targets. But it would seem preferable to let the requirement of making an adequate phonological description of a language constrain the theory of phonetics we must develop so that it will be able to account for the cases we have described above by having more categories or possible combinations of categories than are necessary to satisfy the requirement of providing an adequate interpretation of linguistic items in non-linguistic terms. The latter requirement may, however, result in the theory having some categories which are necessary at the level of systematic phonetics, but which are not required at the phonemic level. Thus it seems probable that at the phonetic level we may need categories distinguishing between syllabic and non-syllabic sounds; but with the possible exception of Japanese (McCawley, personal communication), this distinction is not needed at other phonological levels. Thus in English, the surface contrasts between *hungry* and *Hungary* or between *coddleing* and *codling*, do not occur in the underlying forms.

Targets and Categories

So far, no explicit statement has been made about the relations between segmental targets, and the categories which characterize the segments. Earlier investigators (Holmes, Mattingly and Shearme, 1965; and an earlier draft of this monograph) envisaged a system in which segments such as phonemes or allophones were assigned values by finding the entry corresponding to each segment in a look-up table. Recently Kim (1967) has proposed an interesting way of making generative rules for going from the systematic phonetic level to the level of physical phonetics. He was concerned with interpreting systematic phonetic features (or categories) in terms of acoustic parameters; but the underlying ideas are equally applicable to physiological specifications of speech. His principle (slightly reinterpreted) is to begin by a rule assigning values which would specify a neutral position of the speech mechanism. He then interprets each group of categories, such as those required for vowel height, by means of another rule giving a value for a degree of movement and a set of rules such as:

(a) if high up two degrees
(b) if low down two degrees
(c) if tense to a certain degree (of tenseness) then up to that same degree (but of movement).

Kim's rules are expressed in a more elegant way, using a notation with variables which is partially of his own devising. But even the formulation given above enables us to appreciate a number of the advantages of his system. Firstly it clearly distinguishes between the properties of the language and those of the individual speaker. The rules for assigning values to specify the neutral position, and the rules that give the magnitude of a degree, correspond to properties of individual speakers. Changing them will be equivalent to changing the personal quality of the speaker being synthesized. But the rules concerning the interpretation of categories, such as those exemplified above, are properties of the language. A second advantage is that this system may be more economical in that fewer rules may be needed for the interpretation of categories than for the interpretation of segments; this, however, has not been shown conclusively in that it appears that the rules may have to be very context restricted. The third advantage is that it has often been shown (cf. Halle 1964; Chomsky and Halle 1965) that phonological rules are expressed most appropriately in terms of features (our categories) rather than segments.

Intrinsic and Extrinsic Allophones

In whatever way it is arranged, the table of values will have to contain sufficient information to distinguish between all the linguistic oppositions within a language. The most well known of these involve categories at the phonemic level. There is no need to discuss here the concept of a phoneme (or morphoneme--I follow Chomsky (1964) in finding no necessity for two separate levels, and will normally say phoneme and phonemic where older American linguists might have said morpho(pho)-neme and morpho(pho)nemic). Similarly the contrastive nature of suprasegmental units such as pitch and stress is sufficiently well known to require no further elaboration. But there are other contrasts which cannot be specified simply by taking into account the segmental phonemes, and suprasegmentals such as pitch and stress. Consider the example (suggested by Bloch in a personal communication) *I'm going to get my lamb prepared* as opposed to *I'm going to get my lamp repaired* in which the sequences of segmental phonemes and stresses are identical. At some point in the description of English these two will be distinguished by a juncture mark, or word boundary, or similar device. But later this distinction will have to be expressed in terms of categories characterizing the allophones of the particular segmental phonemes, since only these units (and the suprasegmentals which are irrelevant here) can be interpreted in physical terms. Allophones such as p and p^h will have to be composed of categories with separate listings in the table of values for English. Their parametric specifications are different; and, although the differences are linguistically predictable in terms of juncture, they cannot be produced by means of conjoining rules.

Junctures cannot be listed in a table of values as having certain relative formant frequencies and durations; they can be taken into account only by noting their effect on other units. Allophones which are generated in this way (or through the effect of other higher level units such as stress or vowel harmony marks) may be called extrinsic allophones, in contrast with those which are due to the partial overlapping of the articulations of adjacent phonemes, which will be called intrinsic allophones. These two kinds of allophones have to be distinguished because categories characterizing extrinsic allophones have to be given individual listings in a table of values, but those for intrinsic allophones do not.

Some further examples may help in clarifying the differences between these two types of allophones. We have already mentioned the well known fact that in the English words *key* and *car* the initial consonants differ in that the stop in the first word has a more forward articulation than that in the second. These two stops are intrinsic allophones, since, as we have seen, it is possible to postulate a single ideal position for English initial k and a rule which enables one to calculate the actual articulatory position by taking into account the overlap with the articulations required for the neighbouring vowels. Similarly the difference between the voiced r in *dry* and the largely voiceless r in *try* can be predicted by a rule specifying the nature of the overlap between the states of the glottis in the adjacent sounds. None of these allophones has to be specified separately in a table of values. But the difference between the t in *top* and the t in *mountain* (which in many forms of American English may be accompanied by a glottal stop) cannot be ascribed simply to the overlapping articulations of neighbouring sounds; nor can the difference between the r in American English *reed* and that in *deer*; nor the differences such as those in timing between the two n sounds in *nun* or the two k sounds in *kick*. These extrinsic allophones must be characterized by categories with different listings in the table of values.

Both the ideal positions in a table of values and the conjoining rules for specifying intrinsic allophones are language dependent. There are many linguistic universals; but, for example, the effect of neighbouring vowels on the articulation of velar stops is not one of them. This may be seen by comparing English and French. In both languages the initial stops vary in much the same way in pairs such as English *key* - *car* and French *qui* - *car*; but there is a much greater difference between the final stops in French *pique* - *pâques* than there is between those in English *peak* - *pock*. Some of this difference may be due to the differences between the vowels in the two languages. But an explanation of this kind is not sufficient. It seems that in English coarticulation consists mainly of anticipation of the following item, but in French preceding vowels have as much effect on a consonant as following ones. In other words, the conjoining rules for English and French have to be different.

The table of relative values which is being suggested here as a necessary part of the description of each language must be able to generate a far greater number of allophones than the number of phonemes in the language. This is the price we have to pay for the fact that in all languages phonemes combine to form larger phonological units such as stress groups and words. We may find it profitable to reinterpret Sweet's (1890) division of phonology into analysis and synthesis. In the description of a language there is a section of statements or rules which allow us to convert simultaneous strings of items such as tone groups, stress groups, word boundaries, and phonemes into a single string of allophones specified in terms of phonetic categories; this part of the description involves phonological analysis. Then, as a new technical term, we may say that synthesis is the equally necessary part of the description formed by the rules which convert these allophones into observable speech.*

Differences Between Languages

Sounds which can be correlated with phonemic oppositions or extrinsic allophones within a single language clearly need to be specified in different ways. But before we can formalize a set of categories we must state a principle for classifying similar sounds which never contrast within a single language, such as Hindi dɦ and English ð. We can obviously reduce the number of categories required for describing the languages of the world by classifying similar phenomena in different languages as variants belonging to the same category. But unless we are careful this is going to lead us into the game playing arbitrariness of some adherents of Jakobsonian distinctive feature theory. There are two principles which will help us here. In the first place

*It would probably be advisable to alter various parts of the preceding sections to emphasize the difference between rules of the phonology proper, and what we may now call interpretation conventions which might be a better name for the "rules" involved in synthesis. No linguistic claims are made by the form of the interpretation conventions.

we should try to account for linguistic oppositions in terms of categories which are necessary and sufficient, in the sense that they fit into a scheme of speech synthesis which will account for all and only the linguistic aspects of speech. We should not be satisfied with categories that are merely possible in that they characterize some and not all of the distinctive aspects of the linguistic oppositions. In the second place we should describe sounds in terms of the same categories only if they differ in degree and not in kind. This second principle can be put more formally (and in a more widely applicable form) by saying that *two different phenomena should be described in terms of the same category if and only if the category has measureable properties and the phenomena can be differentiated simply in terms of numbers specifying the degrees in which they possess these properties.*

We will find it useful that this second principle applies just as much to the description of two different sounds within a language as to sounds in different languages. It makes it immediately obvious that, despite the phonetics of Trager and Smith (1951), at the systematic phonetic stage of the description the first sound in English *hot* and the last sound in English *law* can no more be in the same categories than the first and last sounds in English *hang*. As there is no gain in grouping any of these items together at an earlier stage in the rules (except a pseudo-economy which later has to be resolved), we never have any motivation for this part of the Bloch-Trager-Smith analysis of English. Our knowledge of phonetics and speech synthesis by rule is still too limited for us to be sure of the details, but I would guess that a principle of the kind italicized above should enable us to evaluate many other disputed points in the phonology of English and other languages.

Competence and Performance

Chomsky (1965) has proposed that we should distinguish between a speaker's competence, which is his knowledge of the rules of his language, and his performance which is what he actually does when speaking. It is not easy to say how these terms fit into the model proposed here. Quite clearly, if we consider physical phonetics necessary for the testing and verification of a linguistic description, then the rules that go from the phonetic level to physical phonetics are rules for the interpretation of a model of a speaker's competence. But equally obviously there is no reason to believe that speakers organize their performance in terms of rules of this kind. We do not know how the intention of saying something develops into the formation of a sentence with an appropriate syntactic structure. But everything we do know about the performance of skilled muscular movements such as those involved in speech suggests that they are not arranged in terms of segmental features, but in terms of larger units. This point has been made elsewhere (Ladefoged 1967; cf. also Fromkin 1966); but an additional example may help. In order to produce correct utterances speakers or speech synthesizers have to interpret the category bilabial in different ways when it occurs in different contexts. Thus the lip gesture involved in a bilabial stop is different from that in a bilabial fricative. Furthermore, Fromkin (1966 and personal communications) has shown that the muscular activity is different in a final voiceless bilabial and a final voiced bilabial; and in initial and final bilabials; and in bilabials preceding different vowels; and in bilabials in stressed and unstressed syllables. Hence the interpretation of this category involves a series of complex context restricted rules. It is perfectly possible to organize the controls for a speech synthesizer in this way; an appropriate program is now being developed at UCLA using a small digital computer. A system of language description of this kind has a great deal of intuitive appeal, and is very elegant in that it uses the smallest possible number of primitive categories, and the greatest number of formal rules. But human beings are not necessarily elegant, and do not work like computers. The indications from neurophysiology and psychology are that, instead of storing a small number of primitives and organizing them in terms of a large number of rules, we store a large number of complex items which we manipulate with comparatively simple operations. The central nervous system is like a special kind of computer which has rapid access to the items in a very large memory, but comparatively little ability to process these items when they have been taken out of memory. There is a great deal of evidence that muscular movements are organized in terms of complex, unalterable chunks of at least a quarter of a second in duration (and often much longer) and nothing to indicate organization in terms of short simultaneous segments which require processing with context restricted rules.

It follows from this that the relation between a speaker's competence and his performance may be very complex. It is also possible that a speaker's competence should be related to his performance both as a listener and as a speaker. Studies of the perception of speech are of little assistance in this respect, since almost the only relevant data available to the experimenter are the responses of subjects. But these responses are discrete items which reflect an encoding of the sensations corresponding to the incoming stimuli. Subjects can report only the results of their perception; they cannot make explicit anything about the process of encoding sensations (cf. Popper 1967). It is possible that this process involves something similar to the activity reflected in the rules for going from systematic phonetics to physical phonetics but in reverse. At the moment this hypothesis is uninteresting because it is untestable. Meanwhile these interpretation rules are in any case an essential part of the description of a language, since without them the description has no substance and is itself untestable.

It is tempting to imagine that there might be some universal perceptual categories so that we could test a grammar without going to the length of requiring the interpretation of our categories to include all and only the information necessary for synthesizing natural speech. It would be so much simpler if the interpretation had to indicate only some of the perceptual data which a hearer might be assumed to use in decoding utterances; and it would seem that a legitimate aim for a linguistic description might well be to characterize simply the important perceptual features in terms of a set of universal perceptual categories. There are two objections to this approach. In the first place unless we generate all and only the information for synthesizing speech, we cannot be sure that we have in fact synthesized the most important perceptual features. Speech is enormously redundant, and subjects in experiments can base their responses on cues which they might never use in other circumstances. Secondly, it is difficult to know what could be meant by universal perceptual categories, if perception is taken to include the encoding of sensations. Perception in any other sense is unknowable; and in this sense is dependent on the listener's previous linguistic experience. Subjects in experiments can usually categorize speech sounds only in terms of sounds which they can make themselves. Trained phoneticians may do better, but if we use them as subjects, so that we partially remove the influence of the subject's linguistic background, then we might just as well remove all such influences and devise categories based on the physical and not the perceptual properties of speech sounds. If a particular perceptual category has any universal validity, it is almost certainly because of some property of the physical stimulus. Categories which simply reflect the linguistic competence of listeners cannot be used in a theory of phonetics without destroying the whole concept of phonetic specifiability.

Multi-Valued Categories

We will discuss in a later section the formal relations among phonetic categories. But while we are considering the general nature of phonetic specifiability we will note an apparent conflict between the requirement that a phonetic theory should provide categories which are useful for designating the classes of sounds needed in phonological rules, and the requirement that these categories should be interpretable and have a non-arbitrary relationship with some extralinguistic events. The conflict arises because it has been claimed that the first requirement is most easily met if sounds are classified entirely in terms of categories consisting of binary features (cf. Jakobson 1962, Halle 1964, Chomsky and Halle forthcoming). But in some cases this condition is impossible to maintain. Some linguistic contrasts, such as tones, consist of a number of items arranged along a single continuum. If it is true that the items are distinguished simply by the degree to which they have a given property, and if it is also true that the definitions of features must be expressed in relative terms, it is difficult to see how to deal with this situation by means of binary features. For example, let us consider a language (such as Yoruba) with three contrasting tones, high, mid, and low. We can distinguish between them by using two binary categories, high - non-high, and low - non-low. This is fine if the features are completely abstract, and are being used simply as labels for natural classes; any set of objects can be arbitrarily classified in this way. But it is not possible if the features have to be given properties. High can be defined as possessing a comparatively rapid rate of vibration of the vocal cords;

and non-high as possessing a comparatively slow rate of vibration of the vocal cords. But low also has to be defined as possessing a comparatively slow rate of vibration of the vocal cords; and non-low as possessing a comparatively high rate of vibration of the vocal cords. Given the requirement that the features should be relative, so that high differs from non-high in a relative way and no arbitrary reference points may be used, and given that only one continuum is involved, it is logically impossible to define a distinction between these two features. In writing phonological rules we may want to group the tones by means of binary divisions; and a ternary system may be very inconvenient. But if we are not playing games and indulging in hocus pocus linguistics, and if we want our descriptions in terms of classificatory features to be mappable onto phonetic data, then we cannot logically have two independent relative binary features within one continuum.

The situation is even worse if we are trying to describe four objects distributed along one parameter, as in a tone language such as Tiv; and it is not improved by the use of features such as high - low, and mid-- non-mid, as suggested by Wang (1967). It is impossible to define Wang's terms, given the requirements that features are not ranked relative to one another. Wang does not insist on these phonological prerequisites, and defines his features in terms of an arbitrary median pitch which serves as a reference point. But this median point obviously varies from individual to individual. Nevertheless Wang has to state its value before features such as High-non-high and mid-non-mid can have any kind of substance. This seems to be confusing things which are properties of the language with those which are properties of the individual. One of the great virtues of the descriptive scheme proposed by Jakobson is that it makes it clear that linguistic propositions are dependent on properties which can be given substance in purely relative terms and which are independent of the physical characteristics of individual speakers. Thus compactness and gravity are defined independently of one another and without any reference to the particular mouth shape or cavity resonances of an individual. It would be odd if features of tone were different in this respect.

No binary system reveals the fact that high is related to mid-high in the same way as mid-high is related to mid-low, and in the same way as mid-low is related to low. This kind of relationship is important in many tone lowering rules. And presumably we want a theory of phonetic description and an associated evaluation criterion that enables us to say that a rule lowering each of three tones by one degree is more general than a rule which says that some of these tones change one way and others another. But in a binary system this is not so.

It can be shown that there are similar weaknesses in specifying vowels in binary terms. Put simply, the argument is that if sounds differ in and only in a single property on a given line, then it is impossible to divide them into more than two independent but relative groups. You cannot make a useful second division on the line without knowing where the first one came.

Chomsky and Halle have been foremost among those pointing out that linguistic descriptions must involve a phonetic component which is independent of the language under description. The claim being made here is that phonetic specifiability, when applied to phonological descriptions, should mean that it is possible to generate testable physical events. We must have a theory which allows us to map descriptions onto observable data. And the Jakobsonian theory of binary but relative features does not allow us to do this. Even though the phonetic matrix proposed by Halle (1964a) may now include multi-valued entries, Jakobsonian distinctive features cannot be used to categorize sounds in the underlying classificatory matrix because the matrix they define cannot always be given a phonetic interpretation. But, as Chomsky and Miller (1963) have said: "It is an extremely important and by no means obvious fact that the distinctive features of the classificatory phonemic matrix define categories that correspond closely to those determined by the rows of the phonetic matrices."

Since multi-valued categories seem unavoidable if we are to maintain phonetic specifiability, it seems that we cannot use phonlogical rules which rely on binary categories. In fact it will be shown that there are a great many advantages (and few disadvantages) to a multi-valued system.

<u>Linear Ordering</u>

Items which differ only in the degree in which they possess a given property such as tongue height, pitch, or (perhaps) place of articulation, may be said to be linearly ordered. The existence of sets of items of this kind must be recognized in the theory of phonetics. There are two notions which should be formalized. Firstly we want to state that adjacent items (such as high and mid vowels) are more related and form a more natural class than non-adjacent items (such as high and low vowels); and secondly we want to ensure that the relation between a given item and the next higher up the scale (e.g. mid and high) is exactly the same as the relation between one lower down the scale and the first item (i.e. low and mid). Phonological rules involve both these notions; the necessity of a specification in terms of natural classes has been discussed already; and it will be shown that the notion of direction which is implied by linear ordering is useful not only in descriptions of tonal phenomena, but also in descriptions of vowel changes such as those in the English great vowel shift.

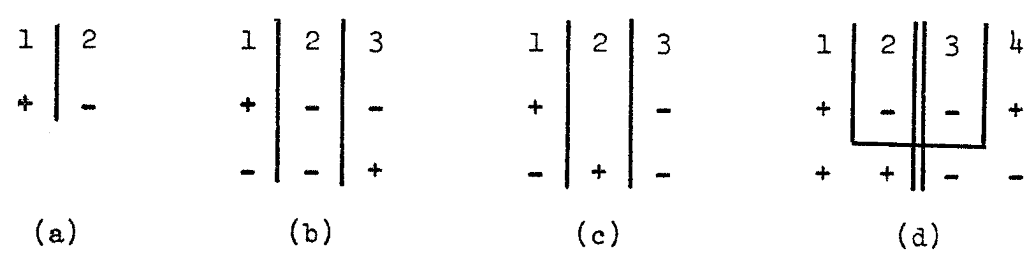


Figure 12: Some relationships between sets of linearly ordered items and binary classifications of these items.

Figure 12 demonstrates some simple points concerning binary categories and linear arrangements of items. If there are only two items, then obviously the desired relations are obtained by using a binary classification. If there are three itmes, there is more than one way of classifying them in binary terms. We can use a system such as that suggested by Halle (1964) for vowel height, in which the three categories traditionally high, mid, and low are classified as being + diffuse, - diffuse and - compact, and + compact. This system, which is represented in Figure 12 (b), shows that the middle item is related to both the first and the last, in that it shares non-membership in a category with each of them; but the first and last items are not related. This may be considered to be a correct statement of the relationships; but it has been achieved in the binary system only at the expense of creating categories which are not truly independent and which have to be defined in such a way that they cannot both have positive values at the same tiem. Futhermore the relation between the first and second categories is not the same as the relation between the second and third.

The system shown in Figure 12 (c) is that used by Wang (1967) for classifying three level tones (high, mid, and low) in terms of the two features High - Non-high, and Central - Non-central. This system has the advantage over that in Figure 12 (b) in that the classification of the middle term may be completed by making it more similar to the first or to the third, whichever seems more in accord with the circumstances in a particular language. But the relation between items 1 and 2 is still not the same as that between items 2 and 3.

If there are four items we can get part of the desired result by using an arrangement of two binary categories such as that used by Jakobson, Fant and Halle (1951) for classifying the places of articulation which occur in p , t , c , k , and that used by Wang (1967) for languages with four level tones. In such a system, which is illustrated in Figure 12 (d) the first and second items, the second and third items, and the third and fourth items are related by sharing membership or non-membership in a category, whereas the first and third items and the second and fourth items are not related in any way. These are some of the relationships which need to be expressed. But this result is achieved only by declaring that the first and fourth items share membership in a category. In other words, Jakobson and his colleagues have to maintain that in *all* languages having the four items p , t , c , k (or the items m , n , ɲ , ŋ) p and k (or m and ŋ) are more closely related than t and k (or n and ɲ); and has to claim that, as a linguistic universal, in all languages with four tones the highest and the lowest form a natural class. This classification is certainly not desirable from the point of view of expressing the relationships of items known to be linearly ordered with respect to a given property. The first and last items might, by coincidence, have some other property in common, but that would have to be shown in each case. In addition, the binary classification does not show that the first, second and third items, and the second, third and fourth items form natural classes whereas the first, third and fourth items and the first, second and fourth do not. And, as in the three item case, there is no expression of the concept of direction, whereby the difference between the first and second items should be identical with that between the second and third items and with that between the third and fourth items.

A simple notation can be used to express the desired relations in a multi-valued system. The items which are linearly ordered with respect to one another (e.g. high, mid and low vowels) may be assigned consecutive integers (e.g. 1, 2, 3) as subscripts to the name of the parameter (e.g. Height, which would be defined in terms of vowel quality). Then we may define a natural class as the sounds specified by any parameter with a single integer subscript or with a set of consecutive integer subscripts, the larger the set the greater the value of the natural class.

As an example of the explanatory power of such a system we may consider the great vowel shift in English. As has been pointed out by Chomsky and Halle (1965), we want to show that there is a relation in present day English, between such words as: *line* - *linear*, *supreme* - *supremacy*, *sane* - *sanity*, and other pairs of words involving similar changes in back vowels (where, however, there are additional complicating changes). Putting this another way, we want to formalize the statements, common in children's spelling books, which apply to large sets of unrelated words; our rules should be able to explain why the schoolbooks say that *i* can be pronounced long as in *mile* or short as in *mill*, *e* can be long as in *theme* or short as in *them*, *a* can be as long as in *rate* or short as in *rat*. Each of these alternations is due, in some sense, to historical sound changes in the pronunciation of English. What we want to know is what all these sound changes have in common. If we can show that there are simple underlying rules, we will have, not a complete explanation of why the sound changes took place, but at least a plausible explanation of why these particular sound changes occurred together.

The major phonetic relationships we want to account for are shown in Table 51.

Table 51. Some systematic correspondences which must be included in descriptions of English phonology.

Orthography	Short Monophthongs	Long diphthongs, beginning	Height of short vowel
i	high front	low central	3
e	mid front	high front	2
a	low front	mid front	1

We may assume, for this argument, that the short forms are more nearly the basic, underlying forms from which the long forms may be derived. Chomsky and Halle (1965) do this using the Jakobsonian binary features and a variable, α (alpha), which is defined as having the value + or the value −. They can then write a rule of the form: α(Feature → −α(Feature), which is a notational shorthand for the two rules (1) +(Feature) → −(Feature) and (2) −(Feature) → +(Feature). Chomsky and Halle have to use two α switching rules (which can be combined into one rule schema by additional notational devices). By the one rule high vowels become mid and mid become high, and by the other mid vowels become low and low become mid. There are two objections to rules formulated in this way: firstly, as Stockwell (1967) has pointed out the switching of sets of vowels is an unlikely process from an historical point of view, and in this sense the rules are not explanatory; secondly, it is apparent from what has been said above that quite disparate processes are claimed in the interchanging of high and mid vowels and the interchanging of mid and low vowels, if these interchanges have to be expressed in binary terms. Although these rules may be descriptively adequate, it is doubtful whether they really explain anything.

A better pair of rules would account for the fact that both the mid and the low vowels behave in the same way (in that they become raised) whereas the high vowels behave differently (in that they become centralized and lowered). If we number the vowel heights so that low = 1, mid = 2, and high = 3, then we can express our rules thus:

	vowel		vowel	
(1)	3 height	→	1 height	
	front		central	
	vowel		vowel	
(2)	x height	→	(x + 1) height	
	front		front	condition: x = 1 or 2

Rules of this kind seem to have far more explanatory power, if this term is considered to have its usual scientific meaning (cf. Brown 1963). An explanation of an event is a statement which adequately describes what happens in a way which illuminates the underlying processes. It should be in terms of previously defined variables, but it is not necessarily the most concise, or most simple, arrangement of symbols.

An interesting possibility might be to combine the rules suggested in the form:

	vowel		vowel	
(2, 3)	x height	→	(x + 1) height	
	front		central	in the environment height 1
			or	
			front	in other environments

Condition: 4 height = 1 height

This condition, which implies that vowel height is a cyclically ordered linear continuum, seems to be unmotivated and have no real explanatory power. It is like the switching rules discussed above, in which an α variable is used. It is ingenious in that it involves a notation which conveniently summarizes other rules. But there is no independent evidence suggesting why linguistic (or other) events could be organized in this way; and therefore such rules cannot be said to explain the processes involved.

Multi-valued Phonemic and Phonetic Characterizations

The relation between systematic phonemic descriptions and systematic phonetic descriptions in a multi-valued system should be made explicit. Let us assume that there is a parameter which may be called vowel height. As was indicated in the previous chapter, it is probable that no language has more than three contrasts on this continuum; and some languages have only two. At the systematic phonemic level, languages may be specified by integer values 1 and 2, or 1, 2, and 3. But at the systematic phonetic level, where we want to characterize the sounds more precisely, a more precise set of values is possible.

We could profitably take over a familiar convention, and write parameter values between slashes (e.g. /1/ or /2/) when we are at the systematic phonemic level; in such circumstances the number of different integers would specify simply the number of contrasts which a given language uses within this parameter. Then at the systematic phonetic level we could put the parameter values between square brackets (e.g. [1·3] or [3·5]). These values are, of course, arbitrary in the sense that the numbers (and names, such as high, mid and low) are assigned arbitrarily to different degrees of the property represented. But when this assignment has been made, it is possible for the phonological rules to generate more precise phonetic values for the phones of a language and to specify the systematic phonetic properties which characterize one language as opposed to another. Examples of the use of these two forms of representation will be given in the next chapter.

3

U N I V E R S A L   P H O N E T I C   P A R A M E T E R S

The oppositions which have to be distinguishable by means of a theory of phonetics have been indicated in the first chapter. In this chapter we will outline the parameters that can be used for this purpose, and indicate the range of values for each parameter. The majority of the parameters are stated in physiological rather than acoustic or auditory terms. In one sense it would be simpler if all of them could be specified in physiological terms; but, as we have seen, we do not yet have a valid set of parameters for specifying tongue positions in vowels. And in another sense it would be more convenient to make all our specifications in acoustic terms; in the present state of our technology, acoustic speech synthesizers are far more convenient for generating the speech sounds of a wide variety of languages by rule. In this book physiological properties have been chosen wherever possible, largely because in this way we achieve a much simpler statement of the possible combinations of parameters.

Glottal Parameters

When we were discussing phonation types in Chapter 1 we listed seven possible states of the glottis, each of which precludes all the others. But although this is an important fact which must be recognized in our theory of phonetics, it is also true that no language uses more than four of these states. Furthermore some of the seven are more like points on a continuum, which it might have been better to split up into a greater number of categories: if we had done this we would have been better able to account for both the additional contrasts, such as those in Korean and Indonesian, which we had to leave out, and the alternations, such as those in Western Popoloc between an extreme form of laryngealization ("creak") and a glottal stop which we had no way of explaining.

Figure 13 illustrates one possible way of considering the parameter which underlies some of these different phonation types.

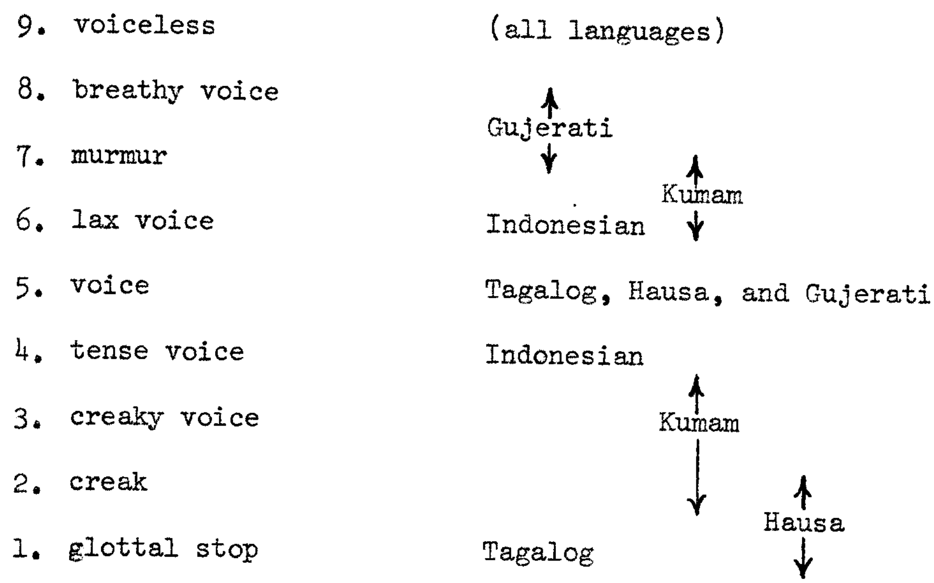


Figure 13. The dimension of glottal constriction. As this dimension is considered to be a continuum, the numbers and labels attached are largely arbitrary. No language contrasts more than three points on this scale.

There is a continuum extending from the most closed position, a glottal stop, through to the most open position observed in speech, which is that in voiceless sounds. Starting from a glottal stop it is possible to pass through a form of laryngealization (here called creak) in which the whole glottis remains constricted except for a small opening in the anterior portion; then through another form, creaky voice, in which a larger proportion of the glottis is vibrating; then, after further releasing the degree of constriction, through stages which we may call tense voice, voice, and lax voice (though of course recognizing, as with all the stages on this continuum, that there is no predeterminable point at which, for instance, tense voice should be considered to become voice). Further relaxation leads to a widening of the glottis, particularly in the posterior portion so that lax voice becomes murmur, in which only the anterior portion is vibrating. This state can, arbitrarily, be distinguished from one in which there is an even greater rate of flow through the glottis which we may now call breathy voice. Finally, when even the anterior portion of the glottis is so far apart that it cannot be set in vibration we have the voiceless position.

No language has contrasts involving more than three states and most languages use only two states within this continuum, which we may call that of *glottal constriction*. This being so, we can usually, in the classificatory matrices describing phonemes, specify this parameter simply in terms of binary possibilities marked /1/ and /0/, and then, in the phonological rules leading to the descriptive matrices, rewrite these items in terms of the appropriate values. But in languages such as Kumam, Western Popoloc, and Gujerati, in which the glottal constriction parameter can have more than two values, we might use the integers /0, 1, 2/ in the classificatory matrix; then as a result of a series of context restricted phonological rules, we may rewrite these values into say, [3, 6, 9] for some Kumam sounds at the systematic phonetic level, or [5, 8, 9] for some Gujerati sounds at this level. As noted before, the particular numbers (and names) used for the different degrees of glottal constriction are the product of an arbitrary assignment within the theory of general phonetics. But once this assignment has been made, the phonological rules for individual languages can generate phones which can be compared not only with each other, but also with the phones of other languages.

There is a great deal of explanatory power in the concept of a parameter of glottal constriction on which some of the glottal states are rank ordered. A parameter of this kind makes a number of linguistic facts easier to explain. Murmured or breathy voiced sounds are in 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 sounds and voiceless sounds, which is what is required in descriptions of Kumam. Furthermore this formulation assists us in making statements about intrinsic allophones. It will be remembered that intrinsic allophones differ from one another in degree rather than in kind. It is now apparent that, for instance, the allophones of /h/ which occur in *That hat* and *My hat* vary in the degree of glottal constriction, and the variations are predictable from the glottal constrictions in the adjacent sounds. The alternations between laryngealization and glottal stop which occur in languages such as Western Popoloc may be explicable in a similar way.

The only major difficulty with this oversimplified description of the states of the glottis is that it leaves no way of accounting for whispered sounds. A number of languages (such as French and Wolof) have contrasts between whispered and voiceless sounds in the environment of pause, which are in complementary distribution with contrasts between voiced and voiceless sounds in other environments. In these circumstances whisper would appear to be an intrinsic allophone of voice. I am not sure of the best way of dealing with this problem.

The other state of the glottis which has been left out of the discussion in this chapter is that which occurs in aspirated sounds. This state might have been regarded as an extension beyond the voiceless position in the glottal constriction parameter, in that it designates sounds in which the minimum degree of glottal constriction is maintained for longer than usual. But it is possible to derive a more appropriate set of natural classes for use in phonological descriptions if it is considered to be part of a separate parameter, to be called *glottal timing* which specifies the moment of onset of regular voicing. This parameter is also a continuum along which we may consider a number of values as shown in Figure 14 (cf. also Figure 2). Most languages use only a binary opposition, and no language contrasts more than three possibilities.

1	voicing throughout	French	English (1–2)	Thai	voiced
2	voicing in part				partly voiced
3	voicing starts immediately after	French (3–4)		Thai	voiceless unaspirated
4	voicing starts shortly after		English (4–5)		slightly aspirated
5	voicing starts considerably later			Thai	aspirated

Figure 14. The dimension of glottal timing. As this dimension is considered to be a continuumthe divisions are arbitrary. No language contrasts more than three points on this scale; and it applies only to stops and fricatives.

At first glance it might seem as if there is a great deal of overlap between the two glottal dimensions; and it is certainly true that there is some redundancy in specifications using both parameters in that, for example, a glottal stop necessarily implies no vibration. But many of the glottal contrictions can be regarded as occurring with several of the glottal timing possibilities. The independence of the two parameters is demonstrated by the data shown in Figure 15, some of which should be regarded as tentative.

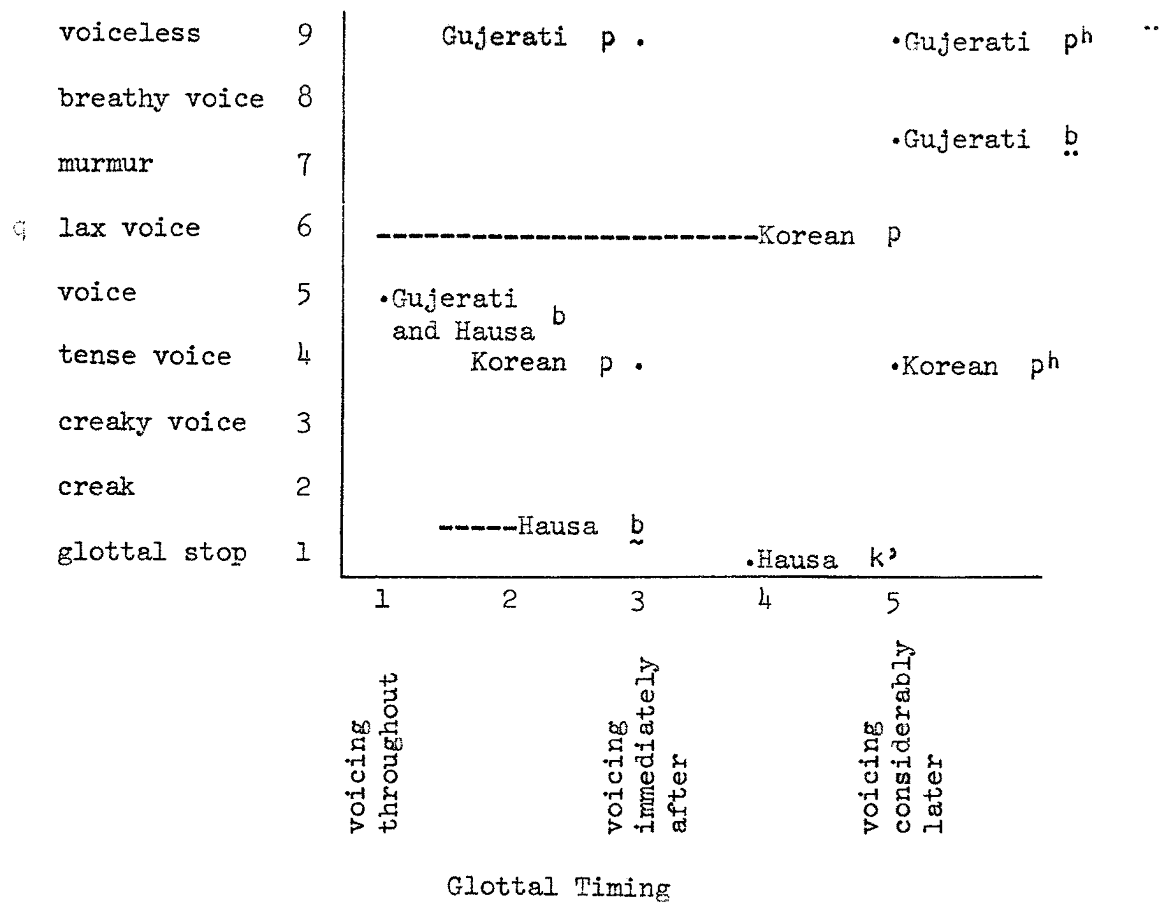


Figure 15. A (tentative) arrangement of some stop consonants, showing the relation between the two glottal parameters.

It would seem that this kind of analysis differentiates sounds in terms of appropriate natural classes. The Korean stops are arranged as suggested by Kim (1966): there are two sets of stops with tense voicing, which are distinguished from each other by variations along the glottal timing parameter; and there is a third set, which is distinguished from the other two by having lax voicing, but for which the glottal timing parameter is not fully specified, since the allophones may vary from being fully voiced to being slightly aspirated. Among the Gujerati stops, the maximum differentiation is between b and p^h ; b̤ and p^h have something in common; and p lies between these two and b .

A suggestion is also made in this figure concerning the categorization of some of the Hausa stops. This language contains a set of voiced stops, a set of voiceless stops, and a set of glottalized stops which includes the laryngealized sounds b̰ and d̰ and the ejective k' . The latter sound (like other ejectives) can be said to have a glottal closure (state 1) throughout the articulation and for a short period after its release, before the beginning of regular voicing. It thus appears on the chart quite close to b̰ and d̰ . At the systematic phonemic level all these sounds may be said to have the maximum degree of glottal constriction, and be marked /1/, Hausa having the possibilities /1, 2, 3/ for the parameter at this level. Through the application of the phonological rules they would come to have the systematic phonetic values indicated in Figure 15.

The relationship between the two glottal parameters must also be remembered in connection with other data concerning airstream mechanisms. In a large number of languages (Swahili, many dialects of Hindi), there is a close connection (perhaps free variation) between fully voiced sounds (glottal timing 1) and voiced implosives. When we take this in conjunction with the necessity of having a natural class of the kind described for Hausa, and the fact that in many languages (including some forms of Hausa) there is free variation between laryngealized stops and implosives, it seems apparent that we must regard sounds with a glottalic airstream mechanism as being on a third parameter which intersects with these two in such a way that the values for glottalic sounds are in the appropriate regions.

The Hausa case can be handled within the Jakobsonian framework by the use of the feature checked (glottalized) - unchecked. But it is often impossible to achieve correct natural classes in other languages by means of binary specifications such as voiced - voiceless, aspirated - unaspirated, or fortis - lenis. This is especially evident if we require these terms to define real properties and not to be names summarizing disparate phenomena occurring in different circumstances.

Nasality

In the first chapter we discussed the possibility of having three basic categories: nasal (all the airstream going out through the nose), nasalized (part going out through the nose, and part through the mouth) and oral (no air going out through the nose); and it was pointed out that this might be an unmotivated complication in the theory in comparison with a two way categorization: nasal (soft palate down) and oral (soft palate up). Neither of these two forms of description will take into account the case of Chinantec and any other languages which have a contrast between oral, lightly nasalized and heavily nasalized vowels; so these data cannot be used for giving preference to one formalization over the other. Nor will either of them enable us to categorize prenasalized stops such as mb nd etc. as units. It might seem as if we could solve this problem by using the three way categorization. Then we could say that mb and nd belonged in the middle category of nasalization which was defined as having part of the air going out through the nose. But in these cases it is really all the air going out through the nose for part of the time, not part of the air going out through the nose throughout the segment as in nasalized vowels. Nasalized vowels and prenasalized stops are qualitatively different, and cannot be specified in terms of a single paramenter without departing from the criteria suggested in the previous chapter.

The solution to the problem of which of the two descriptive systems to use is partly an empirical matter, resting on which gives the best sets of natural classes. For the moment we will prefer the system with the smaller set of primitives, and say that there is parameter of *nasality* which is defined in terms of the degree of velopharyngeal closure. In the vast majority of languages, at the systematic phonemic level only two degrees need be assigned: if the soft palate is raised so that there is virtually complete closure, the sound may be said to have a value of /0/ on this parameter (i.e. to be oral); if it is not it will have a value of /1/ (and be called nasal). In a few languages, such as Chinantec, it may be necessary to assign an in between value, even at the systematic phonemic level; but examination of the underlying forms in these languages may show that they differ in the number of segments involved, so that the three way contrasts between oral, lightly nasalized, and heavily nasalized vowels are really contrasts of the form a - ã - ãn or a - an - ãn the final consonants not appearing in the phonetic output. Variations in the degree of nasality at the systematic phonetic level are, of course, common in many (perhaps most) languages. An additional parameter, *prenasality*, is also necessary. This parameter is defined as being dependent on the degree of velopharyngeal closure which occurs before another articulation such as an oral stop or fricative in circumstances which require the whole complex to be considered as one phonological unit. Only the values 0 or 1 are possible for this parameter.

Airstream Mechanisms

We do not have to distinguish between all four of the possible airstream mechanisms: pulmonic egressive (as in plosives), glottalic ingressive (as in implosives), glottalic egressive (as in ejectives), and velaric ingressive (as in clicks). A pulmonic egressive mechanism occurs in all sounds, so the presence of this mechanism does not have to be marked; but, in cases where the pulmonic mechanism is the only one, we still have to mark the fact that the other airstream mechanisms are not present. These other airstreams form a set of mutually exclusive possibilities (except that the voiced implosive mechanism may be combined with the click mechanism; this rare possiblity will be considered later). The simplest formal method for categorizing segments is therefore to set up a parameter with four states: 0 = no additonal (only pulmonic) airstream; 1 = velaric; 2 = glottalic egressive; 3 = glottalic ingressive. It is difficult to say what claims about the relation between airstream mechanisms would be implicit in such an arrangement; but it is clearly not in accord with the constraint that the members of a category should differ only in degree, and that the differences should be quantifiable in terms of a single measurable parameter. From a physiological or acoustic point of view there is no single factor underlying clicks and sounds produced with other airstream mechanisms.

An alternative specification is to use the fact that implosives and ejectives differ in terms of the single parameter, rate of vertical movement of the larynx towards the lungs. At the systematic phonetic level, ejectives would then be specified by a negative number, and implosives by a positive number, the magnitude of the number indicating the degree of force with which the glottalic airstream mechanism was used. Such a specification provides a nice way of describing weakly implosive allophones which occur in many languages in certain environments.

A suitable name for this parameter is *glottalicness.* We may consider the zero or unmarked member to be that which occurs in ordinary pulmonic sounds in which there is no movement of the larynx. If we do this we can handle the fact that implosive and ejectives are not known to occur at the same place of articulation by making it an empirical observation that at the phonemic level languages never contrast more than two points on this parameter. If a language is found which has both implosives and ejectives at the same place of articulation, it is simply the empirical observation which has to be changed, not the set of parameters which are used for specifying languages.

The presence or absence of a velaric airstream mechanism (or, at the systematic phonetic level, the degree of presence of the mechanism) has to be specified separately, by means of a parameter which we may term velaricness. This is as it should be since it permits the possibility of clicks and implosives co-occurring; and it also allows us to give a correct phonetic description of the difference between the powerful clicks in Nguni languages, and the wide range of weak uses of the velaric airstream mechanism which occur in labial velars in West African languages.

Manners of Articulation

It would be difficult to find a single parameter underlying all the manners of articulation we discussed earlier. But it is equally obvious that these are not all separate, unrelated, phenomena, which have to be described by a set of binary parameters each indicating virtually only presence or absence of an event. We must account for the fact that, for example, fricatives are more related to stops, than stops are to vowels; and that stops may alternate with flaps or taps, and taps alternate with trills. The most appropriate way of handling some of the major manners of articulation is to set up a parameter which may be called *articulatory stricture*, which, like all parameters is a continuum, but which may be said to have three principal values, corresponding to the degrees of articulatory stricture in stops, fricatives and approximants. Different approximants may have very different values of articulatory stricture at the phonetic level; but these will all be irrelevant at the classificatory level unless it turns out that there are grounds for saying that there is a natural class consisting of the vowels i , u , o , ɔ , a , (and similar vowels with different degrees of lip rounding). All of these vowels involve a considerable degree of narrowing of the vocal tract, and may be called near vowels, as opposed to other vowels which may be called far vowels. If we need these two natural classes of vowels, then we may consider this parameter to have four principal values: (1) stops; (2) fricatives; (3) approximants - near vowels; (4) approximants - far vowels.

It should be noted that this parameter does not account for the relation between high, mid and low vowels. As we saw earlier, this relationship may be expressed in terms of a parameter that may be called *auditory height*. We may consider this parameter to have three principal values at the classificatory level: (1) as in low vowels; (2) as in mid vowels; and (3) as in high vowels and semivowels (which will be non-syllabic approximants in our system). Again, if it turns out that extra values are needed (perhaps to account for the vowels of Ngwe in Table 42 and Figure 4), then it is only an empirical observation which needs to be changed, not the set of parameters.

The articulatory stricture and auditory height parameters clearly overlap to some extent. But if we are to follow our principle of having a single variable underlying each parameter, these two cannot be combined, since in the one case the variable is the cross-sectional area of the vocal tract at the point of maximum constriction, and in the other it is an auditory quality most nearly ascribable to the frequency of the first formant.

So far, we have not specified a parameter which will account for the relations between stops, flaps, taps and trills. We can do this by specifying a parameter of *rate*. Most sounds can be said to have a normal (or 0) value on this scale. But when an articulation such as a stop occurs more rapidly, it may be specified by a lower value. The flapped sounds which occur in Hausa and in some Indian languages can be assigned a value of less than 0, say -1, on this arbitrary scale. The slightly more rapid movement which occurs in a tap as in Spanish *pero* may be assigned a value of -2. More rapid movements (caused not by muscular activity, but by the airflow), as in the Spanish trill in *perro*, may be assigned a value of -3. A parameter of this kind seems to capture the correct relationship between taps and trills. These sounds sometimes contrast (as in Spanish); but they more often alternate or are in free variation (as, for example, in Hausa and some forms of Scottish English).

Nevertheless this characterization may lead us into some difficulties, since, if we postulate a parameter of this kind, we cannot avoid using it to describe other contrasts between sounds which are due to differences in articulatory rate. Thus the difference between vowels and semivowels must be described in this way, semivowels being given a negative value as for a flap or a tap. But what is far more important is that all contrasts in length must be placed on this parameter. Sounds which are some arbitrary percentage longer than normal may be assigned a value of 1; and higher values may be assigned to sounds which are even longer as discussed in the first chapter. This is a useful generalization of a parameter; and it is insightful in that it shows that semivowels, taps, and flaps cannot be long. But it also suggests that trills cannot be long; which may not be a correct observation.

In classificatory matrices the value of this parameter need not be specified for the majority of sounds, but following the proposal of Chomsky and Halle (forthcoming) it can be designated as unmarked. A universal redundancy rule will then be all that is needed to supply the appropriate value in systematic phonetic matrices. The role of redundancy rules, and the evaluation of phonological descriptions of particular languages will not be considered further here, partly because these topics have been so much more ably discussed by others (Chomsky and Halle, forthcoming, Stanley 1967), and partly because the main emphasis in this monograph is simply on developing a theory which will provide adequate phonetic descriptions while being compatible with the general requirements of a generative phonology. But although there is little consideration of such phonological problems in all cases it should be completely clear how the concepts presented here make it possible to specify categories and how they make it possible to interpret these categories in terms of values of specifiable parameters.

The two parameters we have discussed will account for most manners of articulation. I am not at the moment certain whether it would be advisable to set up an additional parameter for differentiating types of fricatives. Distinctions such as those between grooved and slit articulations involve articulatory differences which are not specifiable in terms of the articulatory stricture or the articulatory rate parameters. They are probably best accounted for in terms of the parameters for describing tongue shape which will be given later.

The status of affricates must also be considered. It seems that we must have a special parameter for specifying the degree of *affrication* of stops. Only two values contrast at the phonemic level; but at the systematic phonetic level there are characteristic differences between languages requiring intermediate values as in other parameters.

An additional parameter is needed to distinguish between central and lateral articulations.

It is difficult to conceive what might be meant by in-between values in the case of this parameter, which we may call *laterality*. At the systematic phonemic level sounds are either lateral or they are not (in which case they are central); and the same seems to be true at the phonetic level.

Places of Articulation

The main reason for considering all sounds, including both vowels and consonants, in terms of the same set of places of articulation is that this allows us to give a good explanatory account of many phenomena involving assimilations. The phonological dsecriptions of languages frequently require statements showing how s is realized as ʃ in the environment of high front vowels, and how vowels become diphthongal with a w offglide in the environment of velar consonants. We could, of course, simply make statement of this kind and not go any further. But if we did just this we would not be saying very much about how languages work. Our descriptions would contain equally conplex statements for becoming ʃ in the environment of i as for s becoming ʃ in the environment of o , despite the fact that these are clearly not equally complex phenomena. We can write rules which have much more explanatory power if we consider vowels such as i to be similar to post-alveolar sounds, u to be similar to velar sounds, and perhaps o and ɔ to be in the uvular and pharyngal regions respectively (cf. Peterson and Shoup 1966).

Bearing this in mind, and neglecting for the moment the actions of the lips, we may say that there is a parameter of *articulatory place* which we may designate in terms of seven phonetic categories: (1) dental, (2) alveolar, (3) post-alveolar, (4) palatal, (5) velar, (6) uvular, (7) pharyngal. All these terms are simply labels for arbitrarily specified points on the continuum formed by the roof of the mouth and the back wall of the pharynx. As we implied by the empirical observations in Chapter 1, probably no language requires the specification of more than five of these points at the classificatory level.

We must note also that this parameter may not be relevant for some segments. This is not the same as saying that it can be left unmarked, in the way discussed before in relation to the parameter of rate. In the latter case, the designation of a segment as unmarked meant that it could be considered to have a normal rate, not that it did not have a specifiable articulatory rate (which would be an impossibility). But if we consider the actions of the lips separately (a procedure which will be justified later) then for segments such as p the tongue may have no specific position; in a sequence such as that in *apt* it will move from the target for æ to the target for t , without having a target associated with p . A similar situation occurs for segments which we wish to describe simply in terms of the state of the glottis, such as many examples of h and ʔ . It is impossible to specify a target position for the tongue or the lips in these sounds; and when no place of articulation is specified, the articulatory stricture and auditory height parameters are also unspecifiable.

Other parameters may also be considered in this way. The parameter of velaricness is obviously sometimes not relevant, rather than being specifiable in terms of its degree of presence or absence. The degree of possession of this parameter refers to the degree of suction created by the movement of the back of the tongue; a zero value would mean that there was no suction created (as in Late k͡p , cf. Ladefoged 1964) rather than implying that the velaric airstream mechanism was not able to operate. But in the majority of sounds the prerequisite lingual gesture does not occur.

Many other parameters have restrictions on their relevancy. The prerequisites for prenasality occur only with stops and fricatives; for auditory height only with approximants; and for affrication only with stops. Our description of segments in terms of parameters will therefore sometimes contain blanks, which have a different formal status from numerical values, or specifications of values as marked or unmarked. Blanks which occur at the classificatory level will remain blank even at the systematic phonetic level, since they denote parameters which, for the particular segment being described, are always irrelevant. (A similar point is made by Wilson (1966) in a discussion of the Jakobsonian distinctive feature system.)

If we describe the parameter of articulatory place in the way suggested above, we will not be able to account for all the actions of the tongue which occur in speech. We need another parameter, *apicality*, in order to be able to specify the difference between apical and laminal articulations. This parameter is irrelevant when neither the tip nor the blade of the tongue is involved. On the classificatory level only two values are needed. At the systematic phonetic level the value 0 may be assigned to sounds made with the absolute tip of the tongue, and the value 1 to articulations involving an arbitrary location considered to be maximally far back on the blade of the tongue. If it is found necessary to specify sounds made with the underpart of the blade (as, for instance, in the phonetic characterization of the extremely retroflex sounds which occur in some Indo-Aryan languages) then negative values could be assigned.

We need one more parameter of tongue shape in order to be able to specify the difference between tense and lax vowels. So as to make it quite clear that this parameter is being defined by reference to the action of the tongue alone, it will be called *tongue tension*. The best physical measure underlying this parameter is the degree of longitudinal bunching in the tongue (cf. Ladefoged 1964a). This may be said to be 0 when the tongue is not bunched up, the intrinsic muscles are relaxed, and the shape and position of the body of the tongue are both determined by the extrinsic lingual muscles; and it may be said to be 1 when there is the maximum degree of bunching.

The muscles which are primarily responsible for the bunching up of the tongue are the inferior longitudinals. These muscles are opposed by the superior longitudinals, which are responsible for the curling up of the tip of the tongue, and the hollowing of the upper surface. We could therefore assign negative values to this parameter when some degree of hollowing or retroflexion occurred. At the moment I am not sure how to regard the relation between this parameter and the apicality parameter. They are sometimes interdependent; but they can also be independent. It is possible that between the two of them we can characterize many of the tongue positions in fricatives which were giving us difficulty earlier. But obviously further research is needed.

Our parameters have to be set up so as to provide us with adequate natural classes as well as adequate specifications of target positions. Consequently we must be able to specify the class of sounds with secondary articulations, even though they all involve a sequence of targets. Just as we have included a parameter of affrication because it is needed in phonological rules, so we must also have a parameter, *secondary articulation*. When it is relevant, the specification of a value for this parameter will imply that there is an additional near approximant articulation, the place of which could be defined in terms of the places of articulation given above, but which might be better labelled in terms of four principal phonetic values, corresponding to (1) post-alveolar (if sounds with a secondary articulation of retroflexion can be considered in this continuum), (2) palatalization, (3) velarization, (4) pharyngalization. As far as I know, at the systematic phonemic level, no language contrasts more than two of these possibilities; and in the majority of cases it is often more explanatory to consider a language as simply contrasting one of these possibilities and its absence.

Note that this parameter is like others such as velaricness and affrication which are often not relevant. When a value is assigned to any of these parameters it implies that the prerequisites for the parameter are met. In the case of the secondary articulation parameter, this means that there is a near approximant articulation. Leaving this parameter blank means that there is no such articulation. Consequently in evaluating the proposed parameter system with respect to the number of states that it will generate, we must consider blanks as indicating a possible state. Even if there were no languages which have contrasts involving kinds of secondary articulations, we would still have to regard the secondary articulation parameter (and the velaricness and affrication parameters) as determining two states at the phonemic level. Similar observations have often been made about the Jakobsonian feature system. Some features, such as stridency and vocalicness, determine the presence or absence of an acoustic or articulatory condition; others such as gravity and diffuseness specify one or other end of the scale of the quality constituting the feature. But in the case of this binary system, each feature contributes two states, irrespective of whether it is of the one kind or the other.

We should also note that the number of features or parameters, and the number of states which each of them determines are not the only things which have to be taken into account in evaluating a theory of phonetics. Quite apart from all the considerations concerning the prerequisites of the theory (as discussed in Chapter 2) and the explanatory power of the parameters (which is largely an empirical matter, justifiable by reference to the phonological description of individual languages) there remain questions concerning the economy of the feature or parameter set. A set of, say, 15 binary features can be less economical than a set of 20 multi-valued parameters, if there are clearly statable dependencies among the parameters, and less obvious relationships in the set of binary features. In the present edition of this monograph there will be no attempt at calculating the number of possible combinations of parameter values at the phonemic level; but it seems probable that it would be a more reasonable number than that generated by the Jakobsonian feature system.

We must now set up parameters which will acount for different labial gestures. The actions of the lips are considered separately, because they can (and often do) act independently of the tongue. If we tried to set up a single parameter for all possible places of articulation it would have to include the double articulations labial velar, labial palatal, and labial alveolar; and then it would be impossible to speak of this parameter in terms of a single physical continuum. Instead we will postulate two parameters to be associated with actions of the lips: *labiality*, which will have two principal values, dependent on whether the lower lip is articulating with (1) the upper lip (for a bilabial segment) or (2) the upper teeth (for a labiodental); and *lip rounding* which will have arbitrary values between -1 (lips spread), 0 (neutral) and 1 (close rounding). In the majority of segments, when there is no specific target of either kind for the lips, both these parameters are irrelevant, and may be left blank. When the labiality parameter is specified a parameter of *labial stricture* must also be specified. This, like the parameter of articulatory stricture, may be said to have three principal values, corresponding to the degree of articulatory stricture in stops, fricatives, and approximants. Note that through having the two parameters labiality and lip rounding we are able to account for the differences between sounds with labial compression (which may be bilabial approximants, with or without a simultaneous lingual target) and these with lip rounding (which may be closely rounded sounds, with or without a simultaneous lingual target).

In the majority of languages the parameters of labiality and articulatory place are never both relevant in the specification of a single segment. As a result, the parameters of articulatory stricture and labial stricture need not be specified separately. Instead, a convention can be adopted whereby the value of the stricture which is specified is taken to apply to the relevant parameter.

The parameters that have been proposed here are sufficient to take care of almost every articulatory gesture. But they may not provide all the natural classes that are required. Some additional possibilities may be necessary because speech sounds can be grouped not only in articulatory terms but also in auditory terms. We have already set up one auditory parameter, auditory height, which is required for providing the appropriate relations among sounds. Very possibly another would be useful in order to account for the similarity between velar and labial voiceless fricatives (which underlies the sound change from x to f which occurred in English words such as *rough*) and between velar and labial voiceless stops (as discussed by Jakobson, Fant, and Halle, 1952). Note that if we did postulate such a parameter, which, following Jakobson, we could call *gravity* we would not have to imply that bilabial and velar nasals were more closely related than alveolar and velar nasals, just because this kind of relationship is true for voiceless stops and fricatives. We are suggesting the possibility of an additional parameter with definable auditory or acoustic characteristics -- pitch or frequency of the turbulent noise. Where this is not present or inaudible, the parameter would obviously be irrelevant.

Subglottal Pressure and Syllabicity

In the first chapter we observed that an increase in subglottal pressure occurred in certain consonants (typically in the so-called fortis consonants) and during the whole of a stressed syllable. Fortis consonants are comparatively rare, and I do not know enough about the phonology of any of the languages involved to be able to say whether there are any reasons for setting up more than one parameter to account for them and for variations in stress. For the moment we will postulate a single parameter, *pressure* which will be assigned a value of 0 when the subglottal pressure is the normal appropriate for the utterance, and a higher value say 2, when it is at its maximum value. I would guess that we do not need more than three contrasting values at the most abstract phonological level.

It seems as if we do need separate parameters for syllabicity and stress, since there is no single measurable phenomenon underlying these two features. It would have been nice if it had been true that syllables were accompanied by a pulse produced by the muscles of the chest, and stressed syllables were simply chest pulses reinforced by the abdominal muscles (Stetson, 1951). But it is not (Ladefoged 1958, 1967). Accordingly we must say that there is a parameter of *syllabicity*, the underlying characteristics of which are at the moment obscure, but which is probably related to the timing and compounding of articulatory gestures. Syllabicity either occurs or does not occur. It cannot be said to be a continuum, and is thus not like other parameters. (It is, indeed, the most ill-defined and unsatisfactory of all the parameters that have been suggested here; much research is still needed in this area.)

Tone

The last parameter we have to consider is the one about which I know least. For the moment we will postulate a single parameter of *tone* to take account of all pitch phenomena. But in all probability this parameter should be split in some way, so as to account for the natural classes that are required in descriptions of tonal phenomena, as suggested by Wang (1967). We have already noted that Wang's solution is not entirely appropriate, partly because he relies entirely on binary categories and does not allow the possibility of tones being related in a linearly ordered series, and partly because it is not always possible to give a physical interpretation to his features. But it must also be admitted that he has a more insightful view of tonal phenomena than any I can suggest here.

Summary of the Parameter System

A complete summary of the proposed system is given in Table 52.

Table 52. Phonological Parameters

	Name of parameter	Maximum Contrasts		Arbitrarily specified phonetic points
1.	Glottal constriction	3	1	glottal stop
			2	creak
			3	creaky voice
			4	tense voice
			5	voice
			6	lax voice
			7	murmur
			8	breathy voice
			9	voiceless
2.	Glottal timing	3	1	vibrating at start
			2	vibrating in part
			3	vibration starts immediately after
			4	vibration starts shortly after
			5	vibration starts considerably later
3.	Nasality	2?	0	oral (velic closure)
			1	nasal (Maximal velic opening)
4.	Prenasality	2	0	not prenasalized
			1	prenasalized
5.	Glottalicness	2	-1	Ejective (glottis moving air upward)
			0	Pulmonic
			+1	Implosive (glottis moving air downward)
6.	Velaricness	2	0	no click
			1	click (maximum ingressive velaric airstream)
7.	Articulatory stricture	3	1	stop
			2	fricative
			3	approximant
8.	Labial stricture	3	1	stop
			2	fricative
			3	approximant
9.	Auditory height	3	1	low
			2	mid
			3	high
10.	Rate	3?	-3	trill
			-2	tap
			-1	flap
			0	normal
			1	long
			2	extra long

Table 52. Phonological Parameters (continued)

	Name of parameter	Maximum Contrasts		Arbitrarily specified phonetic points
11.	Affrication	2	0	unaffricated
			1	affricated
12.	Laterality	2	0	central
			1	lateral
13.	Articulatory place	5	1	dental
			2	alveolar
			3	post alveolar
			4	palatal
			5	velar
			6	uvular
			7	pharyngal
14.	Apicality	2	0	tip of tongue
			1	blade of tongue
15.	Tongue tension	2?	-1	tongue hollowed
			0	
			1	tongue bunched
16.	Labilality	2	1	bilabial
			2	labiodental
17.	Lip rounding	2	-1	lips spread
			0	lips neutral
			1	lips closely rounded
18.	Secondary articulation	2	1	retroflexion
			2	palatalization
			3	velarization
			4	pharyngalization
19.	Pressure	3?	0	background subglottal pressure
			1	slightly increased subglottal pressure (fortis, or stressed)
			2	maximum subglottal pressure (strongly stressed)
20.	Syllabicity	2	0	non-syllabic
			1	syllabic (correlates undefined)
21.	Tone	6?	1	lowest pitch
			2	
			3	
			4	
			5	
			6	highest pitch

This set of parameters will obviously be compared with the Jakobsonian distinctive feature system, in which there are 14 or 15 binary segmental features, and the system proposed by Chomsky and Halle (forthcoming) in which there are 27 segmental features, and 12 suprasegmental features. It should be noted that if there are 27 binary features, then there are 54 terms available for classifying sounds at the systematic phonemic level.

The advantages of the system proposed here are: (1) it accounts for a far wider range of data; (2) it is slightly more economical in that only 48 possibilities are postulated for segmental features at the systematic phonemic level; (3) the universal restrictions on combinations of parameters are easier to state, while it remains equally possible to use the concepts of marked and unmarked values for parameters; (4) the system proposed here enables us to write better phonological descriptions of languages (better in the sense that they are usually simpler -- involving a smaller or less redundant set of symbolizations -- and are often much more explanatory); (5) it is based on an explicit set of prerequisites which make it quite clear when, for example, two different parameters must be postulated; and (6) it enables us to go in an explicit way from the highest phonological level of description of a language (containing only statements in terms of the oppositions which occur) to a testable systematic phonetic description (specifying all and only the aspects of the speech sounds which are characteristic of that language in opposition to others).

BIBLIOGRAPHY

Abercrombie, D. (1967) *Elements of General Phonetics* Chicago: Aldine Publishing Company.

Bell, A. M. (1867) *Visible Speech* London.

Bolinger, D. L. (1958) A theory of pitch-accent in English *Word* 14, 109-149.

Brown, R. (1963) *Explanation in Social Science* Chicago: Aldine Publishing Company.

Catford, J. C. (1964) Phonation types in *In Honour of Daniel Jones* (ed. Abercrombie et al.) London: Longmans 26-37.

Chao, Y-R. (1920) A system of tone-letters *Le Maître Phonétique* 45. 24-27

Chinebuah, I. K. (1962) *A Phonetic and Phonological Study of the Nominal Piece in Nzima* unpublished M.A. thesis: University of London. [cited in Abercrombie (1967)].

Chomsky, A. N. (1964) Current issues in linguistic theory *The Structure of Language* (ed. J. Fodor and J. Katz) Englewood Cliffs: Prentice-Hall Inc. 50-118.

-------------- (1965) *Aspects of the Theory of Syntax* Cambridge, Massachusetts: M.I.T. Press.

Chomsky, A. N. and Halle, M. (1965) Some controversial questions in phonological theory *Jour. of Linguistics* 1.2. 97-138

Chomsky, A. N. and Halle, M. (forthcoming) *The Sound Pattern of English*

Chomsky, A. N. and Miller, G. (1963) Introduction to the formal analysis of natural languages. *Handbook of Mathematical Psychology* (ed. Luce, R. D., Bush, R. R. and Galanter, E.) New York: John Wiley and Sons Inc. 3. 269-323

Conklin, H. C. (1949) Preliminary report on field work on the islands of Mindors and Palawan, Philippines *Amer. Anthropologist* 51. 268-73

Doke, C. M. (1931) *A Comparative Study in Shona Phonetics* Johannesburg: University of the Witwatersrand.

Dunstan, Elizabeth (1964) Towards a phonology of Ngwe *Jour. W. Afr. Lang.* 1.1. 39-42

Ferguson, C. A. (1961) (with Moukhter Ani and others) *Damascus Arabic* Washington: Center for Applied Linguistics.

Firth, J. R. (1957) *Papers in Linguistics, 1934-1951* London: OUP

Fischer-Jorgensen, Eli (1952) The phonetic basis for identification of phonemic elements *Jour. Acoust. Soc. of America* 24. 611-17.

-------------- (1960) *Almen Fonetik* Copenhagen: Rosenkilde og Bagger.

Fortune, G. (1966) Unpublished duplicated material on Ndebele.

Fromkin, Victoria (1966) Neuro-muscular specification of linguistic units *Language and Speech* 9.3. 170-199.

Gleason, H. A. (1961) *Introduction to Descriptive Linguistics* (Revised Edition) New York: Holt.

Gleason, H. A. (1961a) Review of "African language studies I" *Language* 37.2 294-308.

Guthrie, M. (1948) *The Classification of the Bantu Languages* New York: OUP.

Halle, M. (1964) Phonology and generative grammar *Structure of Language* (ed. J. Katz and J. Fodor) Englewood Cliffs: Prentice-Hall Inc. 334-352.

-------------- (1964a) On the bases of Phonology *Structure of Language* (ed. J. Katz and J. Fodor) Englewood Cliffs: Prentice-Hall Inc. 324-333.

Halliday, M. A. K. (1963) The tones of English *Archivum Linguisticum* 15.1 1-28.

Heffner, R-M. S. (1950) *General Phonetics* Madison, Wisconsin: University of Wisc. Press.

Hockett, C. F. (1955) *A Manual of Phonology IJAL* 1.24 pt 1. Baltimore: Indiana University.

Holmes, J. N., Mattingly, I. G. and Shearme, J. N. (1964) Speech synthesis by rule *Language and Speech* 7.3. 127.

Hudgins, C. V. and Stetson, R. H. (1935) Voicing of consonants by depression of the larynx *Arch. néerl. Phon. Expér.* 11. 1-28.

International Phonetic Association (1949) *Principles* (Revised edition) London: University of London.

Jakobson, R. (1962) *Selected Writings I: Phonological Studies* The Hague: Mouton.

Jakobson, R., Fant, G. and Halle, M. (1951) *Preliminaries to Speech Analysis* Tech. Report - No. 13, Acoustics Laboratories, M.I.T. Cambridge Mass.: M.I.T. Press.

Jakobson, R. and Halle, M. (1964) Tenseness and laxness *In Honour of Daniel Jones* (ed. D. Abercrombie et al.) London: Longmans 96-101.

-------------- (1956) *Fundamentals of Language* The Hague: Mouton.

Johnson, Jeannette (1966) *Tamazight Phonology* unpublished paper, UCLA.

Jones, D. (1956) *An Outline of English Phonetics* (8th edition) Cambridge, England: Heffer.

Jones, S. (1929) Radiography and Pronunciation *Brit. J. Radiol.* 2.15. 149-56.

Joos, M., ed. (1957) *Readings in Linguistics* Washington: Amer. Council of Learned Societies.

Kelly, John (1966) A note on lip-rounding *Le Maître Phonétique* 125. 8-9.

Kim, C-W. (1966) On the autonomy of the tensity feature in stop classification. To appear in *Word*.

-------------- (1967) *The Linguistic Specification of Speech* Working Papers in Phonetics 5. UCLA.

Kozhevnikov, V. A. and Chistovich, L. A. et al. (1965) *Speech: Articulation and Perception* Moscow-Leningrad, 1965. Translation JPRS 30, 543. Department of Commerce.

Ladefoged, P. (1958) Syllables and stress *Miscellanea Phonetica* 3, 1-14.

-------------- (1962) *The Nature of Vowel Quality* Coimbra: Laboratoria da Phon. Exp.

-------------- (1962a) *Elements of Acoustic Phonetics* Chicago: University of Chicago Press.

Ladefoged, P. (1963) Some physiological parameters in speech *Language and Speech* 6. 109-119.

-------------- (1964) *A Phonetic Study of West African Languages* London: OUP.

-------------- (1964a) Some possibilities in speech synthesis *Language and Speech* 7. 205-214.

-------------- (1967) *Three Aspects of Experimental Phonetics* London: OUP.

Langendoen, D. T. (1964) Review of "Studies in linguistic analysis" *Language* 40.2 305-321.

Lanham, L. W. (1964) The proliferation and extension of Bantu phonemic systems influenced by Bushman and Hottentot *Proc. Ninth Int. Int. Cong. Linguistics* (ed. H. Lunt) The Hague: Mouton 382-390.

Lehiste, Ilse (1964) Acoustical characteristics of selected English consonants *IJAL* 30 Pt. IV.

Lehiste, Ilse and Peterson G. (1961) Some basic considerations in the analysis of intonation *Jour. Acoust. Soc. of America* 33. 419.

Lindblom, B. (1963) Spectrographic study of vowel reduction *Jour. Acoust. Soc. of America* 35. 1773-1781.

Lisker, L. and Abramson, A. S. (1964) A cross-language study of voicing in initial stops: acoustical measurements *Word* 20. 384-422.

Longacre, R. E. (1952) Five phonemic pitch levels in Trique *Acta Linguistica* 7. 62-82.

Malmberg, B. (1956) Distinctive features of Swedish vowels, some instrumental and structural data *For Roman Jakobson* (ed. M. Halle) The Hague: Mouton 316-321.

Malmberg, B. (1963) *Phonetics* New York: Dover.

Moll, K. (1966) (Personal communication).

Merrifield, W. R. (1963) Palantla Chinantec syllable types *Anthropological Linguistics* 5.5, 1-16.

Moore, P. and Von Leden, H. (1958) Dynamic variations of the vibratory pattern of the normal larynx *Folia Phoniatrica* 10. 205-238.

Ohman, S. (1964) Numerical model for co-articulation, using a computer simulated vocal tract *Jour. Acoust. Soc. of America* 36. 1038.

-------------- (1967) Numerical model of co-articulation *Jour. Acoust. Soc. of America* 41.2 310-320.

Pandit, P. B. (1957) Nasalisation, aspiration and murmur in Gujerati *Indian Ling.* 17. 165-172.

Peterson, G. and Shoup J. (1966) A physiological theory of Phonetics *Jour. Speech and Hearing Res.* 9.1. 6-67.

Pike, E. V. (1963) *Dictation Exercises in Phonetics* Santa Ana: S.I.L.

Pike, K. L. (1943) *Phonetics* Ann Arbor: Univ. of Michigan Press.

-------------- (1946) *The Intonation of American English* Ann Arbor: Univ. of Michigan Press.

-------------- (1948) *Tone Languages* Ann Arbor: Univ. of Michigan.

Popper, R. (1967) Unpublished paper UCLA.

Schachter, P. (1961) Phonetic similarity in tonemic analysis *Language* 37.2. 231-239.

Stanley, R. (1967) Redundancy rules in phonology (to appear in *Language*).

Stetson, R. H. (1951) *Motor Phonetics* (2nd edition) Amsterdam: North-Holland Publishing Company.

Stetson

Stevens, K. and House, A. (1955) Development of a quantitative description of vowel articulation *Jour. Acoust. Soc. of America* 27. 484-493.

Stockwell, R. (1959) Structural dialectology: a proposal *American Speech* 34.4 258-268.

-------------- (1967) Mirrors in the history of English pronunciation (to appear in the *Festschrift for Rudolf Willard* University of Texas.

Sweet, H. (1890) *A Primer of Phonetics* Oxford: Clarendon.

Trager, G. L. and Smith H. L. (1951) *An Outline of English Structure* Studies in Linguistics. Occasional papers, 3. Reprinted 1957, Washington D.C.

Tucker, A. N. (1940) *The Eastern Sudanic Languages* London: OUP.

Wang, W. (1967) Phonological features of tone *Int. Jour. Amer. Linguistics* 33.2 93-105.

Welmers, W. (1959) Tonemics, morphotonemics and tonal morphemes *General Linguistics* 4. 1-9.

Westermann, D. and Ward, Ida (1933) *Practical Phonetics for Students of African Languages* London: OUP.

Whiteley, W. H. and Muli, M. G. (1962) *Practical Introduction to Kamba* London: OUP.

Wilson, R. D. (1966) A criticism of distinctive features *Jour. of Linguistics* 2.2 195-206.

INDEX OF LANGUAGES CITED

(Note: I have no first hand data for any of the languages in parentheses.)

Akan languages, 36
Aklanon, 36
Amharic, 17
Arabic, 26, 39
Assamese, 10
Ateso, 12

Bantu languages, 10, 14, 20
50, 61
Bengali, 10
Berber languages, 39
Bihari, 10
Bini, 32, 33
Bura, 12, 28, 29
Burmese, 8, 9, 35
(Bushman), 20

Chadic languages, 12
Chaga, 33, 35
Chinese, 43, 50
Chinantec, 79
Czech, 32

Dagbani, 28
Dravidian languages, 24, 30

English, 5, 7, 9, 15, 19, 26
30, 33, 35, 38, 46, 50, 52
53, 61, 64, 65, 71, 77
English, American, 39, 47, 48
English, Scottish, 9, 46, 81
English, South African, 32
Ewe, 23

French, 7, 14, 26, 35, 42, 43,
48, 64, 76, 77

Gaelic, 9, 33
Gujerati, 8, 9, 10, 75, 76, 77

Hausa, 12, 33, 75, 77, 78, 81
Haya, 33
Hindi, 7, 10, 64, 78
(Hottentot), 20

Icelandic, 9
Idoma, 26
Igbo, 40, 43
Indonesian, 14, 74, 75
Indo-Aryan languages, 33, 84
Italian, 50

Japanese, 61
Javanese, 14
Jinphoˀ, 17, 19

Kamba, 26, 39, 50
Kannada, 30
Khoisan languages, 20
Kom, 39
Korean, 14, 16, 17, 74, 77, 78
Kumam, 75, 76
Kutep, 28, 39

Lango, 11, 12
Late, 83
Luganda, 16, 17

Makua, 61
Malayalam, 24, 25, 33, 35
Mandarin, 43, 52
Marathi, 10
Margi, i, 11, 12, 26, 35
Marwari, i, 10, 19

Navaho, 12, 36, 61
Ndebele, 10, 11, 12
Nguni languages, 20, 80
Ngwe, 35, 40, 41, 42, 43-48,
81
Ngwo, 24
Nilotic languages, 12
Nzima, 28

Polish, 38
Popoloc, Western, 74, 75, 76
Punjabi, 76

Quechua, 26, 61

Russian, 38

Saurashtri, 10
Shona, 10, 26, 28, 35, 39, 76
Sindhi, 10, 17, 18
Slavic languages, 38
Spanish, 33, 81
Swahili, 19, 78
Swedish, 28, 42, 43, 48

Tagalog, 12, 13, 75
Tamazight, 39
Tamil, 33
Telugu, 30
Thai, 8, 9, 77
(Ticuna), 52
Tiu, 13, 14, 61, 68
(Tojolabal), 17
Tsonga, 10
Twi, 36, 50, 61
Tzotzil, 19

Welsh, 35
Wolof, 14, 76

Xhosa, 20

Yoruba, 5, 13, 14, 50, 67

Zulu, 10, 16, 20, 35, 36

GENERAL INDEX

Acoustic parameters, 45
Affricates, 59, 82, 84
Airstream mechanisms, 15, 16, 79
Allophones, 63
Alveolar, 23, 24
Alveolar ridge, 22
Alveolar-velar fricative, 28
Analysis and synthesis, 64
Apical, 23, 24
Apicality, 84
Approximants, 28, 48, 81
Articulatory place, 83
Articulatory process, 22
Articulatory stricture, 81
Aspirated fricatives, 8, 9
Aspirated stops, 8
Aspiration, 6, 7
Assimilations, 83
Auditory height, 81, 84

Bilabial, 22
Bilabial fricatives, 23
Bilabial trill, 35
Binary features, 67
Breathy voice, 6, 9, 75

Cardinal vowels, 41
Checked, 78
Chest pulses, 87
Clicks, 16, 18, 20, 79
Competence, 65
Contour tone languages, 50
Contours, 52
Creak, 75
Creaky voice, 6, 12, 75

Dentals, 23, 24
Depressor stops, 11
Discrete level languages, 50
Double articulations, 26

Ejectives, 16, 17, 78, 79
Emphatic consonants, 39
Epiglottal friction, 39
Extrinsic allophones, 62, 63

Flap, 32
Formants, 45
Fortis, 15, 17, 87
Fricatives, 9, 14, 30, 81
Fricatives, ejective, 17

Glottal constriction, 74, 75
Glottal parameters, 74
Glottal stop, 6, 12, 13, 75
Glottal timing, 76
Glottalic airstream, 16, 17, 79
Glottalicness, 80
Glottalized stops, 78
Gravity, 86
Grooved articulations, 30, 82

Horizontal lip rounding, 38

Implosives, 16, 17, 78, 79
Inner rounding, 38
Interdentals, 24
Intonation, 5, 50
Intrinsic allophones, 62, 63

Labial stricture, 86
Labial-alveolar, 26, 28
Labial-palatal, 26
Labial-velars, 20, 26, 80
Labiality, 86
Labialization, 26, 36, 61
Labiodental flap, 35
Labiodentalization, 39
Labiodentals, 22
Laminal, 23
Laryngealization, 6, 12, 36
Laryngealized semivowels, 11
Laryngealized stops, 11
Laryngealized vowels, 11
Laterality, 82
Laterals, 14, 35
Laterals, voiceless, 6, 7
Lax, 45
Lax voice, 75
Lax vowels, 84
Length, 50, 82

Lenis, 15, 16
Linear ordering, 69
Lingualabial stop, 22, 39
Lip compression, 38
Lip rounding, 86

Manners of articulation, 28
Marked, 84
Morphoneme, 63
Multi-valued categories, 67, 73
Murmur, 6, 8, 9, 10, 11, 12, 18, 20, 75

Narrow vowels, 43
Nasal, 7, 14
Nasality, 78
Nasalization, 13, 14, 36
Nasalized semivowels, 13
Nasals, voiceless, 6
Natural classes, 57

Oro-nasal process, 14, 28
Outer rounding, 38
Overlapping articulations, 38

Palatal, 24, 25
Palatalization, 32, 38, 61
Perceptual categories, 66
Performance, 65
Personal quality, 62
Pharyngal, 26
Pharyngalization, 38
Phonation process, 3
Phonation types, 74
Phoneme, 63
Phonemics, 55
Physical phonetics, 57
Pitch, 5, 50, 52, 63
Places of articulation, 22, 83
Post-alveolar, 24
Preaspiration, 9
Prenasality, 79, 84
Prenasalized stops, 13, 61, 79
Prepalatal, 23, 24
Pressure, 87
Pulmonic airstream, 15, 16, 79

Rate, 81
Register tone languages, 50
Relevancy, 84
Retroflex, 24
Rill spirants, 30

Secondary articulation, 36, 78
Segments, 58
Semivowels, 9, 14, 48
Slit articulations, 82
Speech synthesizer, 58
Stress, 15, 53, 63, 87
Stops, 81
Subglottal pressure, 15, 53, 87
Suprasegmental units, 63
Syllabic sounds, 61
Syllabicity, 87
Syllables, 48
Systematic phonetics, 55

Tap, 32
Targets, 58, 61, 62
Tense, 47
Tense voice, 75
Tense vowels, 84
Terraced level languages, 50
Tight phonation, 9
Tone, 5, 10, 12, 50, 87
Tongue tension, 84
Trill, 32

Unaspirated, 7
Uvulars, 26

Velar, 26
Velar laterals, 36
Velaric airstream, 16, 20
Velaricness, 83
Velarization, 38
Velic closure, 14
Vertical lip rounding, 38
Vocal fry, 12
Voice, 5, 6, 75
Voiceless, 5, 6, 75
Voiceless approximants, 11
Voiceless laterals, 8 35
Voiceless nasals, 8
Voiceless semi-vowels, 9
Voiceless unaspirated stops, 8
Voiceless vowels, 9
Vowel continuum, 45
Vowels, 14, 41, 81
Vowels, voiceless, 9
Vowel-shift, 69, 71

Whisper, 6, 12
Wide vowels, 43