Representing Phonetic Structure

Peter Ladefoged

Hyper feature	Major node	Terminal Feature	[named term]	Traditional term
Place	Labial	Labial	[protruded]	*Rounded*
			[compressed]	*Neutral*
			[retracted]	*Labiodental*
	Coronal	Apicality	[laminal]	*Laminal*
			[apical]	*Apical*
			[sublaminal]	*Retroflex*
		Anteriority	[dental]	*Dental*
			[alveolar]	*Alveolar*
			[postalveolar]	*Postalveolar*
	Dorsal	Back	[front]	*Front* *Palatal*
			[central]	*Central*
			[back]	*Back*
		High	[high]	*High vowel / Velar*
			[mid]	*Mid vowel / Uvular*
			[low]	*Low vowel / Pharyngeal*
	Radical	ATR	[+ ATR]	*Advanced Tongue Root*
			[- ATR]	*Epiglottal*
Stricture		Aperture	[stop]	*Stop*
			[fricative]	*Fricative*
			[approximant]	*Approximant*
		Trill	[+ trill]	*Trill*
			[- trill]	
		Lateral	[+ lateral]	*Lateral*
			[- lateral]	*Central*
Nasality		Nasal	[+ nasal]	*Nasal*
			[- nasal]	*Oral*
Laryngeal		Voice	[+ voice]	*Voiced*
			[- voice]	*Voiceless*
		Glottal aperture	[closed]	*Glottal stop*
			[creaky]	*Creaky / Laryngealized*
			[modal]	*(Modal) Voice*
			[breathy]	*Breathy / Murmur*
			[spread]	*Voiceless*
		Aspiration	[+ aspiration]	*Aspirated*
			[- aspiration]	*Unaspirated*
		Pitch -	(tone and intonational features - not considered here)	
Airstream -	Pulmonic	Pulmonic	[+fortis]	*Fortis*
			[- fortis]	*Lenis*
	Glottalic	Glottalic	[implosive]	*Implosive*
			[ejective]	*Ejective*
	Velaric	Velaric	[+ click]	*Click*
			[- click]	

UCLA Working Papers in Phonetics73 April 1989

Representing Phonetic Structure

Acknowledgements

Many of the members of the UCLA Phonetic Lab have offered me a great deal of help with this book. I would like to acknowledge particularly the contributions of Ian Maddieson, Pat Keating, Bruce Hayes, Abby Cohn and Jenny Ladefoged.

Preface

The plan of this book is as follows. Chapter 1 covers some introductory matters that I consider necessary pre-requisites for a theory of linguistic phonetics. Most of these I have discussed before in the papers mentioned below. Those who know (or do not care about) my thoughts on these matters can skip to the next chapter. Chapter 2 begins with a short discussion of the different kinds of features that are needed to account for observed phonological patterns. The main point of this section is that some patterns can be associated with auditory properties of sounds, and others with articulatory properties. A set of auditory/acoustic features is then defined. This chapter also includes a discussion of why a phonological feature should not be defined in terms of more than one property, and why more than two values of a given property should be permissible for a single feature. The third chapter presents a hierarchical set of features with articulatory definitions. A major aim of this set of features is to account, in itself, for all human articulatory capabilities. The fourth chapter, which is not yet even in a good enough draft form to circulate, will be an account of the kinds of properties that are involved in the production of speech.

Some of the early part of this book are based on enlarged and much revised versions of some recent papers, including especially: "Revising the International Phonetic Alphabet." *Proceedings of the Eleventh International Congress of Phonetic Sciences,* Tallin, USSR. 1987; "Updating the Theory." *Journal of the International Phonetic Association.* 17.1, pp10-14.1987; "Hierarchical features of the International Phonetic Alphabet." *Proceedings of the Berkeley Linguistic Society.* 14 1988; and "The many interfaces between phonetics and phonology." Proceedings of the *Sixth International Phonology Conference,* Krems, Austria. 1988. It also reflects some of the views expressed in older papers, such as "'Out of chaos comes order', Physical, biological, and structural patterns in phonetics." *Proceedings of the Tenth International Congress of Phonetic Sciences,* eds. M.P.R. Van den Broeke and A. Cohen. Dordrecht, Holland: Foris Publications. Pp. 83-95. 1984, "Redefining the scope of phonology." *On Language, Rhetorica Phonologica Syntactica: A Festschrift for Robert P. Stockwell from his Friends and Colleagues,* ed. Caroline Duncan Rose and Theo Vennemann. London: Routledge. Pp. 212-220. 1988. Part of the paper, Mona Lindau and Peter Ladefoged "Variability of feature specifications," in *Invariance and Variability in Speech Processes* , ed. Joseph S Perkell and Dennis Klatt. Hillsdale, New Jersey: Lawrence Erlbaum. Pp464-478, is also incorporated, but with a different conclusion.

1. What do we want to represent?

1.1 Delimiting the scope of phonology.

Before we discuss the nature of phonetic representation, we must consider what it is we want to represent, and why we want to represent it. Phonetics is the study of the sounds of speech, so presumably it is these sounds that we want to represent. But what is a speech sound? This in turn leads to another question, namely, what is speech?

An obvious answer is that speech is the substance of spoken languages, making phonetics the physical embodiment of phonology. But this is only part of the story. Speech has many functions; and not all of them are part of language. Similarly language has many functions; and not all of them are part of speech. When we study phonetic representation we are concerned with more than the relation between phonetics and phonology. This relationship is a major part of what we will be discussing. But we must always bear in mind that there is more to speech than we consider in linguistics, and that in studying the relation between phonetics and phonology our first task is to define the scope of phonology, a task which inevitably leads us to consider what we include within the scope of language.

One of the major functions of language that is not part of speech is to act as a mirror for the world. Language provides us with a way of organizing our experiences, and thus gives us ways of grouping these experiences into categories, so that we can qualify and relate one experience to anther. As a result we are able to form concepts and manipulate ideas. Some kinds of thinking may be possible without language, but we could never develop scientific theories without words. As Wittgenstein (1958) puts it, "The limits of my words are the limits of my world." Our language acts as a model of the world as we know it, much as a map serves as a model for a certain locality, enabling us to plan journeys. This function of language has little or no direct effect on phonology - though it does, obviously, have a considerable effect on how we view semantics; and, as we see, it will also help us to delimit the scope of phonology.

Of much greater importance for our view of phonology is the fact that one of the functions of both speech and language is the passing on of information (or misinformation) that we have about the world. In order to convey this information, meaningful units have to be distinguished from one another. Of course, when we consider language simply as a model of our experiences, the items that represent these experiences have to be distinguished from one another as well. But when language is functioning in this way it does not matter how the distinctions are achieved. Language acts as our mirror of the nature of things, irrespective of whether the words are written, spoken, or simply mental images. We do not need to speak the words of our thoughts.

When we do speak, the ways in which we distinguish words clearly involve phonological units. But speech conveys much more than the meanings of the words, the objective information of the kind used in modeling the world. Many of the subtle distinctions that can be made among sounds are used by some of the other functions of speech. Whenever we talk, we convey not only information about the topic under discussion, but also information about the sociolinguistic group to which we belong. This information is rigidly codified, although often not in terms of discrete oppositions of the kind used in phonological oppositions. It is difficult to say exactly what degree of diphthongization in the vowel in 'plate' marks a person as belonging to a particular social class in London (or Australia, or anywhere else that uses this distinguishing characteristic). But anyone familiar with the regional accents in question can easily place a speaker by these sociolinguistic marks.

Facts of this kind give rise to two questions, one concerned with the formal nature of our theory of phonology, the other being more an empirical issue. The formal question is whether we want to consider sociolinguistic information to be part of phonology. The empirical question is whether sociolinguistic information is conveyed by the same aspects of speech sounds as those that are used for distinguishing the linguistic oppositions among words. To take the latter question first, it is, of course, obvious that speech systems often use the same devices for linguistic and sociolinguistic purposes. In the South of England the words 'cat' and 'castle' are pronounced with different vowels, whereas in most forms of American English the vowels in these two words belong to the same phonological unit. This kind of difference, involving the different distribution of items that might be contrastive, is a frequently used sociolinguistic device, as has been noted by many authors. Thus Gershwin and Gershwin (1933) describe two speakers who have the same set of possible phonemic contrasts but use them in different words, as in 'You say ['iðər] and

I say ['aiðə]. You say ['niðər] and I say ['naiðə].' (It should be noted, however, that Gershwin and Gershwin's phonetic observations are not always reliable. They correctly observe that some people pronounce the word 'tomato' as [tə'meɪtoʊ], whereas others say [tə'mɑtoʊ]; but they further claim to have observed the word 'potato' pronounced as [pə'tɑtoʊ]. This seems very unlikely.)

In addition to these gross differences, there are many other differences between dialects and between the sounds that characterize the accents of different languages, which depend on much more subtle phonetic effects. For example, in a survey of speakers of different types of English, I found that 27 out of 30 speakers of Californian English used an inter-dental [θ] in which the tip of the tongue was protruded between the teeth in words such as "think, thin," whereas 27 out of 30 speakers of different forms of Southern British English used a dental [θ] without tongue protrusion in these words. The difference between a dental and inter-dental [θ] is unlikely to be phonemically contrastive in any language, although it is a reliable marker of a difference in regional accent.

A great deal of evidence has been accumulating recently that demonstrates that languages and dialects are often differentiated from one another in ways that are *not* used to distinguish oppositions within any single language. Ladefoged and Maddieson (1986) have described many small differences between languages that are not used within languages. Similar points with respect to differences among fricatives have been made by Nartey (1982) and by Ladefoged and Wu (1984). Cross-linguistic differences in phonation types that characterize different languages have been described by Lindau (1984). Disner (1980) has described cross-linguistic differences among vowels. Keating (1988) has shown that there are cross-linguistic differences in the extent to which different languages coarticulate adjacent stops and vowels. These and many other papers suggest that the sociolinguistic functions of speech *cannot* be described entirely in terms of the same features as those that are used for describing phonological oppositions.

Nevertheless, this is what some feature theories attempt to do, becoming continually more complicated as a result. Thus Jakobson and Halle (1956) can express more phonetic detail than Jackobson, Fant and Halle (1951); Chomsky and Halle (1968) add still more features; and Halle and Stevens (1971) add complexities so as to be able to describe yet more phonetic differences between languages, replacing the feature Voice by a set of four features, Stiff, Slack, Spread, and Constricted. This enabled them to characterize the phonetic differences between, for example, English [p] as in 'spy' and the Korean so-called lax [p]; but the cost is that they no longer have the more phonologically useful opposition voiced-voiceless, and their feature system is unacceptable for phonological classification (Anderson 1978). Furthermore, all this is done in order to be able to describe how the speech of one group of people differs systematically from that of another group of people. But there is no theoretical or empirical reason to expect speech systems to use the same devices for phonological and sociolinguistic purposes.

There are also other aspects of speech that cannot, and should not, be expressed in terms of any of the traditional sets of phonological features. Another function of speech is to convey the attitude of the speaker to the topic under discussion, to the person addressed, and, indeed, to the world in general. These attitudinal differences may not codified in the same way as linguistic information. They are conveyed largely by intonation, which is, of course, also used for conveying linguistic information. Thus we speak of statement versus question intonation in different languages, and emphatic and non-emphatic statements and questions. But other aspects of intonation, such as sarcastic or simpering intonations are not part of language, and therefore should not be described within phonology.

Similarly, we should not consider emotional effects. There seems to be something in common to expressions of anger, astonishment, sorrow, doubt, and love in many different languages, making them general human traits, rather than linguistic communicative devices. To the extent that they are not the same in all languages, they are learned, cultural, aspects of behavior. Many Englishmen consider it normal to speak in a phlegmatic way with a narrow intonation range that Americans consider as indicative of boredom. The Navaho tend (by American English standards) to speak very softly. As Nihalani (1983) has pointed out, Indian English typically sounds rude or aggressive to speakers of British or American English. All these are learned aspects of particular cultures; but they should not be considered as part of language.

Yet another aspect of speech that should be mentioned here is that conveyed by the style. Again neglecting entirely the lexical and syntactic content, I can say a given sentence in a way that sounds as if I am conveying factual information, or saying a prayer in public, or reading a poem on the radio. These effects will rely largely on variations in the rhythm and intonation. Other cultures have developed additional styles of speech to convey other kinds of situational information. In order to develop an appropriate view of phonology, I will suggest that none of these effects should be considered as part of language.

The final kind of information conveyed by speech that nobody would call a function of language, is that it signals the identity of the speaker. When I walk into a house and call out "Hi, it's me" this is all the information I am conveying. I am not really making a declarative statement. Provided it was normal for me to do so I could have established my identity by saying any other phrase, making it clear that personal information can be distinguished from linguistic information. However, it is sometimes not clear that personal information can be distinguished from socio-linguistic information. We each speak in the way that we do partly because our particular vocal organs have certain characteristics, but also because we choose to use within limits, our own personal style of speech. Often what might seem to be a personal characteristic of a particular speaker is in fact something that he or she has chosen to copy, which is shared by a small sociolinguistic group. But even though this is learned vocal behavior, it is still not part of language.

To be more precise about what should be included within language (and hence within phonology), I suggest that we should consider language to be just the system that we need for modeling our known world. This would be very much like considering spoken language as the direct counterpart of written language. From this point of view it would be inappropriate to speak of a language being *reduced* to writing, implying that some part of spoken language is not present in the written form. It would be better to say that (virtually) all that is *language* can be expressed in speech or in writing — and all the sociolinguistic, attitudinal, emotional, stylistic and personal information, that is left out is not part of what we want to define as language.

The parallel between this view of language and written language is not perfect in that writing may convey some sociolinguistic, personal, or stylistic information. It is impossible to tell from these printed pages whether they have been written by an Englishman or an American, or indeed, by a speaker of one of many other forms of English. You could gain some sociolinguistic information if I were to use certain marked phrases or lexical items such as talking about a 'lift' as opposed to an 'elevator', or a 'full stop' as opposed to a 'period.' In addition, other more informal lexical items might convey stylistic or personal information. But in all these cases, when the written language does convey sociolinguistic or other information, it does so by means of precisely those phonological devices that are used to convey information about the topic under discussion. We do not need a special form of phonetic representation to be able to handle information of this sort. All we need is to be able to identify the linguistic oppositions of exactly the same kinds as those that are used to distinguish words.

A phonetic representation provides part of an interface between phonetics and phonology. As is perhaps obvious already, I do not think that it is proper to speak of *the* interface between phonetics and phonology. There are many such interfaces. One of them can be associated with the kind of phonological theory we would need if we limited ourselves to accounting for the strictly linguistic information. This is like the linguistic information conveyed by a written language with a good orthography (i.e. a written language such as Finnish or Swahili in which there are few letter-to-sound ambiguities such as written English 'read' which can be [rid] or [rɛd], and no sociolinguistic variations, such as British English 'colour' and American English 'color'). It is difficult to define the linguistic information conveyed by such a written language in positive terms, but we could say that it is all the encoded aspects of speech except those that convey information about the speaker's identify, attitude, emotions, sociolinguistic background, or style of speech, in so far as these are not conveyed by syntactic or lexical devices.

Writing also conveys certain aspects of intonation. From the syntax, morphology, word order and punctuation (period, comma, query, quotes, parentheses, italics and space marks) we can determine something (but far from everything) about the intonation that a given sentence could have. As Bolinger (1977) has pointed out, the semantics further circumscribe the possible intonations, but again only to a limited extent. The intonational aspect of regional accents are also not represented on the printed page. When speakers of Irish, Welsh, American, or Scottish English read a page such as this one aloud, there will be differences in their intonation patterns that cannot be ascribed to anything written down. From the point of view we have been developing so far, these differences are not part of the language; they convey only sociolinguistic information about the speaker.

The role of intonation is undoubtedly the most problematic part of this view of language and phonology. Past work on intonation is one of the most extensive bodies of phonological work that is not confined to a spoken equivalent of the written language. (Dialect studies are another.) A generation ago, linguists used to argue whether four pitch levels and a number of junctures were sufficient to capture all the meaningful contrasts in English (see, for example, Trager & Smith 1951, Stockwell 1960). Similar discussions are still in progress using a different set of phonetic devices. Thus Pierrehumbert (1980) has an elaborate discussion of the many different intonations that can occur on the phrase 'Manny came with Anna.' In these discussions the notion of a meaningful contrast is not the same as it is in discussions of other aspects of phonology. It includes aspects of the emotional functions of speech that are conveying the speaker's attitude. Outside of discussion of intonation, phonology is not usually considered to include such things, although they could be aspects of segments, as in lengthening to indicate superlatives [it wəz bɪːːg] [hi wəz grːeɪt]. (A possible exception is Prince (1980), who has a brief discussion of the realization of emphasis in Finnish.) Given the view of phonology we are considering, it would seem appropriate to constrain studies of intonation to those aspects of speech that convey simply linguistic information.

If we limit phonology in this way the relation between phonological and phonetic units becomes much more straightforward. The phonologist is no longer under an obligation to describe the phonetic details that characterize the sounds of one language as opposed to another. This is a slightly different view from that which I have often expressed before (Ladefoged 1971, 1980), and still different from that propounded by Chomsky and Halle (1968), and more recently by Halle (1983, 1988). I used to maintain 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)." (Ladefoged 1971:275.) Chomsky and Halle do not explicitly discuss this point, but their work can be read in a similar way; and a similar point of view has certainly been expressed by Schane (1973) who says "Linguistically significant differences are those which characterize native control of a language." Now, having no reason to believe the distinguishing phonetic characteristics of languages are always expressable in terms of phonological features, and also seeing no formal way in which gradient sociolinguistic features can be considered along with categorical phonological features, I prefer to consider these aspects of speech separately. This view makes the difference between one dialect and another, or one language and another, a part of sociology, describable in the same way as any other indexical behavior, such as the dress, appearance, or patterns of belief that characterize a particular group. Such things are part of culture; and there is good authority for saying: "Linguistics is not to be confused with culture." (Stalin, 1950.) This view of language would also regard phonetic differences conveying emotion, or those aspects of the speaker's attitude that cannot be expressed by syntactic or lexical devices, as part of the subject matter of psychology.

1.2 Universal phonetics and phonology

We will now consider an elaboration of this theory of the proper domain of phonology, which leads to a theory more similar to that of Chomsky and Halle (1968). The major weakness of the theory we have been discussing so far is that it does not allow us to say much about the nature of human language as a whole. To do this properly we must be concerned with the general properties of all languages, not just the set of contrasts within a particular language, which is all that is needed for conveying linguistic information. If we are to say something about the nature of human language we must have an overall phonetic framework that allows us to sum up what is different among languages and what is common to all of them. We want to be able to describe speech sounds in terms of some absolute phonetic standards. Note that this is not the same as saying that we want to include sociolinguistic information within phonology. The reason for setting up absolute phonetic standards is not so that we can compare the physical sounds or articulations of one language with those of another. It is so that we can equate phonological patterns, distinctions, and rules among languages. This kind of phonetic representation will undoubtedly enable us to say more about the phonetic differences between languages. But the phonetic framework will not have been set up for this purpose, and there is no reason to expect it to be fully sufficient in this respect. As we noted above, sociolinguistic differences between dialects and languages are only sometimes the same as differences that distinguish words or in other ways convey linguistic information.

The basic problem in laying out a theory of phonetic representation that will permit discussion of phonological universals is in deciding when a sound in one language is equivalent to a sound in another. This is a historic problem. Before the advent of generative grammar, linguists such as Joos (1950) advocated ad hoc descriptions for each language. Hockett (1955) thought "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 all linguists (including those just cited) in fact use phonetic descriptions that imply some absolute frame of reference. As was pointed out many years ago by Fischer-Jørgensen (1952), 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'.

This problem was first stated even longer ago. We are, in fact, grappling with the first two principles set out by the International Phonetic Association for the establishment of a set of symbols to form an International Phonetic Alphabet. In 1900 the IPA published an *Exposé des principes* containing a table showing the recommended alphabet. This table was set up so that it included "les *sons distinctif* de toutes les langues étudiées jusqu'ici". (My emphasis.) Similarly, the 1912 English version, in a section headed "principles of transcription for languages hitherto not transcribed," notes, long before the phoneme became a popular notion: "It is necessary to ascertain what are the *distinctive* sounds in the language, i.e. those which if confused might conceivably alter the meanings of words." (Emphasis in the original.) The corresponding section in the 1922 *L'Écriture phonétique internationale* uses the then new term 'phoneme' saying: "Pour chaque langue, on représente les *phonémes* ou sons distinctifs, et ceux-là seuls." (Emphasis in the original.) The latest (1949) edition of the *Principles* makes as its first point: "There should be a separate letter for each distinctive sound; that is, for each sound which, being used instead of another, in the same language, can change the meaning of a word. "

The second principle in the current edition of the IPA *Principles* is the one that is directly concerned with equating sounds in different languages. It says: "When any sound is found in several languages, the same sign should be used in all. This applies also to very similar shades of sound." This principle is especially important when taken into account with another IPA practice which has never been formally stated as a principle, perhaps because it is regarded as too obvious to mention. This is the principle that the symbols of the phonetic alphabet should be defined in terms of general phonetic categories very much of the kind that we now regard as features. Phonetic theory in the early days of the IPA was greatly influenced by the work of Sweet and Bell, both of whom had developed systems for classifying all the sounds that were known to be able to distinguish meanings in the world's languages. Bell's *Visible Speech* (1867) and Sweet's *Handbook of Phonetics* (1877) provided iconic symbols for showing the combinations of articulatory elements present in a sound. These same elements (or at least a subset of them) were used to define the symbols of the alphabet. Throughout its history the phonetic alphabet has consisted of symbols defined in terms of intersections of phonetic categories (features). Most of the symbols are defined by the terms naming the rows and columns of the charts, and nowadays by the convention that when there are two items in a single cell the first one designates a voiceless sound (if there is a single item in a cell it is always voiced). In addition a few symbols and several diacritics are defined by supplementary notes. The whole work -- principles, charts, symbols and notes -- constitutes the IPA's theory of phonetic representation.

Given this background we may now compare an IPA description with a feature specification of the kind that is nowadays more common. The location of [m] in the chart explicitly indicates that this sound is:

+ voiced
+ bilabial (1)
+ nasal

By means of the labels along the lefthand edge of the chart, the IPA system also indicates that this is a consonant made with the pulmonic airstream mechanism. In much the same way, Chomsky and Halle (1968:5) note that they will use symbols as "informal abbreviations for certain feature complexes." For them this symbol would be a shorthand way of designating the feature values:

+ voiced
+ nasal
+ anterior (2)
- coronal
+ sonorant
etc.

In both cases several other feature specifications are implied. IPA [m] implies [- dental, - alveolar, etc.; - implosive, -click, etc.]; and in SPE (Chomsky and Halle 1968), it is made clear that there are also a number of other features such as Glottalic the values of which, like some of those noted in (2), can be determined by marking conventions.

It is important at this stage to make a clear distinction between the name of a feature and the possible values that it can have. A feature is a scale; a description of a sound gives a certain value on this scale. I will use certain typographical conventions. The names of features will be capitalized, and a value of a feature will be in square brackets. Thus the IPA feature Vowel Height may be said to have values such as [high], [mid-high], [mid-low] and [low], and the SPE feature High has values [+high] and [-high]. These conventions will be extended later when the hierarchial structure of features is discussed.

Both the IPA and the SPE approaches assume that there is a limited set of phonetic categories; but neither of them states how this set can be delimited. Within general phonetics there is no theory that defines what counts as a speech sound for linguistic purposes. Obviously it is not just any sound that can be made with the vocal organs. We can make all sorts of grunts and groans that could never be considered part of human speech. Accordingly, before we can have a valid theory of phonetics, we must delimit the class of sounds that we wish to represent. We will assume that we know what a language is, and we will put aside the question of deciding whether a given sound is part of a language or not. We will assume that grammarians can tell us what is, and what is not, part of a linguistic system. There are obviously a number of borderline cases. For example in English, is the clicking noise written 'tut tut' or 'tsk tsk' a speech sound, or is it just an exclamation? Presumably the latter, as it is not used in forming any words (one cannot say "he was [ʇʇɪŋ]"). Questions of this sort are outside the province of the phonetician. For us the problem is simply a matter of deciding, before considering any linguistic issues, whether some sound made by the vocal organs could be part of a linguistic system, and [ʇ] could.

As we have noted, in practice phoneticians of all traditions behave as if there were a well-defined framework, that allows them to describe linguistic sounds in terms of what are taken to be extra-linguistic categories such as voicing or nasality. (The precise feature set used is irrelevant to the argument at the moment.) But it should be fairly clear that we do not have a well-defined phonetic framework. Whenever we find a sound that has not been described previously in the phonetic literature, we simply invent new categories, or permit combinations of categories that had previously been declared impossible. Thus when I first heard a labiodental flap in Margi I had no qualms in extending the meaning of the term flap (Ladefoged 1968), despite the fact that the gesture involved is nothing like that in any other flap that had been previously reported. And when I heard strident vowels in !Xóõ, in which the main source of acoustic energy is the movements of the aryepiglottic folds (Traill 1985, Ladefoged and Traill 1980), I simply added ventricular to my list of possible phonation types. Similarly neither I nor any other linguist seems to worry about saying at one moment (Ladefoged 1971) that true velar laterals cannot be made, and shortly afterwards (Ladefoged, Cochran and Disner, 1977) describing these sounds in Melpa and other languages spoken in Papua New Guinea. Nor is it considered odd that in the first edition of my textbook (Ladefoged, 1975) there was no mention of epiglottal sounds, but in the second edition (Ladefoged 1982) they are described along with those made at other places of articulation.

This kind of behavior *should* be worrisome, because it makes it very plain that there is no theoretical basis for the existing phonetic framework. If we can simply add to it or modify it in some way whenever we find a new contrast in an uninvestigated language, then we can hardly say that the framework specifies "the phonetic capabilities of man" (Chomsky and Halle, 1968). All it specifies is the set of contrasts that have been observed in languages to date. In other words, it has a functional linguistic basis rather than an *a priori* basis dependent on our articulatory or auditory capabilities.

An alternative proposal that has been made by Catford (1977a), Lindblom (1984) and others is that the set of speech sounds is, in fact the *complete* set of sounds that can be made with the human vocal apparatus. This seems to be not a very fruitful suggestion. As Pike (1943) pointed out, there are numerous marginal speech sounds — clapping the teeth together, talking with a pulmonic ingressive airstream — so many that it seems profitless to try to classify them. Listing all the possibilities will not help us find out what languages are like. It will simply mean that when we describe a sound in any language in terms of some set of features we will have to invoke a general statement that it also has negative values of a long list of features that no language currently uses and none (we would like our theory to say) could ever use. What we want to know is what is the set of possibilities that could be used.

Another possibility is to argue that there *is* a well determined set of phonetic possibilities and that we are innately endowed with these phonetic capabilities. The problem then becomes one of trying to find out what these innate capabilities are. But this does not seem to be a legitimate position; there is no way of making an independent determination of these capabilities. It is not like saying that there are such things as chemical elements, even though, at a particular moment in history, we might not have discovered the complete set of elements. In the chemical case, if we find something lying around, there are tests that we can do to decide whether it is an element or not. We can predict (if we are as bright as Madame Curie) when we find an apparent gap in the Periodic Table, that there ought to be a substance (which we might call radium) that could be discovered. In the case of speech sounds there is nothing comparable that can be done. There are lots of things available for our inspection — whistling, burping, ventricular phonation — but there is no way of telling whether they are part of the phonetic capabilities of man in a linguistic sense without waiting to observe them in a language. Nor is even this a foolproof test. They might have been phonological features in a language spoken until yesterday in Outer Yucca. Hockett (1955) is correct in saying that there is no principled way in which we can determine the limits of the phonetic framework. As a result, it seems uninteresting to talk about our being innately endowed with a set of phonetic capabilities. This notion is largely untestable. The best that we can do is to test whether the observed phonetic possibilities are in fact things that people use when speaking particular languages.

It therefore seems as if, whether we like it or not, our phonetic framework must be set up simply on the basis of observations of phonological contrasts. These observations will only enable us to say when two sounds in the *same* language are different; they will not help us to decide when two sounds in *different* languages are the same. To deal with this problem we must consider what we mean by being the same. No two sounds in different languages are ever absolutely the same, if only because they are spoken at different times usually by different speakers. For the moment we willl disregard the differences that are associated with the sounds having been produced by different speakers, and imagine that they have been spoken by a perfect bilingual speaker of the languages in question. Our problem is now to decide whether the two sounds are sufficiently different so that, if they had occurred in one language they could have distinguished words. Obviously they must be more than just noticeably different. Considering all sounds that are just noticeably different will result in our having a far larger set of sounds than could be used by languages. Listeners have to be able to distinguish words reliably in comparatively noisy surroundings. Languages tend to lose contrasts that are not easily heard.

Rather than considering all sounds that are just noticeably different from one another it would be better to try to find the set of identifiable sounds. The kind of distinctiveness needed within a language is the possibility of one sound being identified as being distinct from all others even when the other sounds are not present for purposes of comparison. Identifying something means, in most cases, being able to give it a name. But we could think of a sound being correctly identified not only when it has been possible to associate it with a name (or a symbol and diacritics), but also when it has been reliably identified by other means, such as correctly allocating it to a small area of a visual chart. If we were to carry out an experiment to determine the set of possible speech sounds using identification procedures, we would have to use skilled listeners. It would be quite legitimate to do this, bearing in mind that what one person acquires as a skill enabling them to identify given sounds can be acquired equally well by anyone who has the fortune to grow up in a community where the sounds in question are used distinctively.

At the moment we have no way of giving a quanititative account of how far apart two sounds must be for them to be able to be identified separately. Some of the contrasts that languages use seem impossibly small, even to me as a phonetician; but native speakers regard them simply as distinct sounds. Accordingly, when trying to define the set of possible speech sounds we must set criteria so that we do not recognize too many distinct sounds, but nevertheless allow for the fact that languages can use what seem to outsiders to be very subtle distinctions. The work has not been done to enable us to do more than estimate very roughly how many such sounds there are. On the basis of listening tests (Ladefoged 1967) I have previously estimated that it is possible to distinguish about 50 vowels in the plane of the primary cardinal vowels, in which front vowels are unrounded and back vowels are rounded with the degree of lip rounding being predicted from the height. This number could be doubled by adding the possibility of front rounded vowels and back unrounded vowels, together with so-called under-rounded and over-rounded vowels (such as the Assamese low back vowel which has the tongue position of [ɑ] and the close lip rounding normally associated with a vowel such as [u]). With the addition of nasalized vowels (which, even with training and experience are not as distinct as oral vowels), rhoticized (r-colored) vowels, and other possible secondary articulations, the total number of distinguishable voiced monophthongs (to add further constraints) is undoubtedly well above 100. This is all without considering various types of diphthongs that are phonologically single segments with on-glides or off-glides. We must also note possible variations in phonation type. Many languages distinguish sets of slackly voiced (slightly breathy) vowels from regularly voiced (modal) vowels (e.g. Jingpho and other languages of Southeast Asia). Others (such as Mpi, a Sino-Tibetan language) contrast more stiffly voiced (slightly laryngealized) vowels and modal vowels. In calculating the total number of vowels we should consider each vowel as potentially occurring on three different phonation types; it is not at all difficult to distinguish at least this number of different voice qualities.

When we consider all these possibilities it seems that Shaw (1920) may have considerably underestimated the number of distinguishable vowels. In the play *Pygmalion* (and in the *My Fair Lady* version, Lerner and Loewe, 1956) Colonel Pickering expresses admiration for the expert phonetician, Henry Higgins, who is able to distinguish 130 different vowels, as opposed to Pickering's 24. Shaw probably based his estimate of the number of vowels on his reading of Sweet (1877), the acknowledged prototype for Higgins, who provided 72 distinct symbols for vowels without considering diphthongs, differences in phonation types, and other aspects of vowels quality for which he provided diacritics. I would estimate that an expert phonetician could distinguish more than 200 vowels of all types; and sounds that an expert can distinguish would also be clearly distinct for anyone who has been brought up speaking a language using them. No language uses anything like 200 vowels. The largest vowel inventory reported by Maddieson (1984) is in !Xũ, with a total of 45 items consisting of 13 short monophthongs, 11 long monophthongs, and 22 diphthongs. This language does not have a number of clearly distinct vowel types, such as front rounded vowels, or back unrounded vowels, or rhoticized vowels. It is apparent that the limitation in the !Xũ inventory size (to a mere 45 vowels!) is not because the language has exhausted the phonetic possiblities available to it.

The number of distinct consonants is also considerable, even if we limit ourselves fairly strictly to what must be called single segments (i.e. disregarding all affricates, prenasalized stops, etc, although many of them function as single phonological segments). The IPA (1979) chart has 81 symbols for consonants, without taking account of oppositions such as that between dental and alveolar stops (which contrast in many Australian languages), voiceless nasals (as in Burmese), or differences between aspirated and unaspirated obstruents (Sindhi has 25 stop consonants, only 10 of which appear as distinct symbols on the IPA chart). We must also consider all the secondary articulations, such as labialization, palatalization, velarization and pharyngealization, which would far more than double the number of possibilities. And again we have to note differences in phonation type, as well as airstream mechanisms of the kind that form clicks, ejectives and implosives. A very conservative estimate would place the total number of consonantal segments as being up in the hundreds, making the total number of possible contrasting segments as high as 600-800. A comparable number occurs in Maddieson's (1984) survey of the phonological segments that occur in 317 languages, carefully selected so as to exemplify the range of the world's languages. He found that when he considered just the contrasting sounds, without taking into account variations in length, he had to recognize about 650 phonetically distinct segments. Probably no language uses more than about 150, (Maddieson reports !Xũ as having a total of 141, consisting of the 46 vowels mentioned above plus 95 consonants, some of which may better be regarded as sequences.) But a theory of phonetic representation has to consider far more possibilities.

There are many reasons why languages do not use a larger number of segmental oppositions. Perhaps the most important is that they are not necessary; languages can have a sufficiently large stock of morphemes while using only a small number of segmental oppositions. Hawaian has only 13 segments, /p,k,ʔ,m,n,w,l,h,i,e,a,o,u/, but it is a perfectly viable language. (It is being killed by the dominance of another culture, not by any linguistic failings of its own.) The phonological devices used by languages do not require the wealth of phonetic possibilities. Nevertheless these sounds are identifiable and must be taken into account in any theory of phonetic representation. As we will discuss in the next chapter an appropriate way of doing this is in terms of phonological features.

It is convenient to make explicit here one final point concerning the nature of the segments for which a phonetic representation is needed. We will take as a phonetic segment whatever units seem appropriate from a phonological point of view. Pike (1943) spent a considerable amount of time trying to elaborate an algorithm for delimiting phonetic segments. But the ultimate product of his labor is simply to show that no algorithm will work in all cases. His notions of a peak or valley of stricture cannot be used to separate vowels from semi-vowels in words such as [kwɪk] 'quick,' and [kjut] 'cute'; nor will they correctly segment the consonants in [tɹi] 'tree', and [čɪp] 'chip'. However, these problems are of little concern if the phonetic structure is regarded as being determined by the phonology, instead of the other way round, as in Pike's view. The phonetic framework we are concerned with is delimited by the needs of the phonology.

1.3 Language as a social institution

In this section I want to consider why phonetic representation is not concerned with mental images. We have discussed what a language is for, and the limits of what we want to consider as language, but have said very little about what a language is. One popular answer, deriving nowadays largely from the many works of Noam Chomsky, is that a language is a set of rules in a speaker's mind. This is a valid view, accounting neatly for a speaker's ability to generate an infinite number of well-formed sentences from a comparatively small set of primitives. But it is not entirely satisfactory in that it takes two or more to communicate. In most cases (especially when we are considering language as part of speech) there must be a speaker and at least one listener. And the speaker and the listener do not share a mind.

It seems to me that the mental nature of language has been somewhat misleadingly presented by Chomsky (1975). His notion that language is an organ of the mind is not very helpful. It is somewhat like saying that digestion is an organ of the body. Digestion is an ability that involves many components, including some things that are normally called organs such as the liver and the pancreas, as well as a number of other things such as saliva, mastication, and bowel movements. Digestion is like language in that it is a system. But neither of them is an organ in the usual sense.

A better way of describing a language is to consider it as an observable social institution, without having to consider what goes on in people's minds. When we consider any social institution we find that it is governed by different principles from those that govern the behavior of individuals. Principles such as communicative efficiency and identification with a group apply to descriptions of what people do. But a language considered as a system where everything hangs together depends not only on individual acts but also on internal forces in that it is to some extent self-organizing.

In order to make this point clear it is worth considering two other examples of self-organizing social institutions. We may begin by comparing a language with a moral code – a system of value judgments applicable in a given community. Any moral code is clearly a product of a society, and is strongly influenced by the surrounding culture. Moral judgements that originally had some utilitarian function rapidly become ritualized. Like pronunciations of words they are as they are because that is the way things are done in a certain society. But morality is also a property of an individual, at least to the extent that the individual can choose to perform moral acts. Morality, or at least the capacity for performing moral acts, may even be like language, in being innate. Certainly one way to think of morality is an over-developed herd instinct -- a self-organized, innate drive for the preservation of the herd rather than the individual.

The moral code that we observe (or feel guilty about) is only one example of a social institution. As another, very different, example consider the economic system. There are obvious market forces affecting the price of goods and the cost of labor (the far from inexorable 'laws' of supply and demand). There are also Galbraithian forces such as the conflict between the company management (whose aim is usually growth, which leads to bigger managerial responsibility and salaries) and the company ownership (the shareholders) who want bigger profits, which may well be achieved without growth and with less management. All these forces, and many more (government, international affairs, and perhaps morality) add up to form a social institution, the economic system, which nobody understands and which is certainly not part of anybody's competence. Without people there would be no economic system. It is like language in that it takes at least two to trade. Furthermore, just as people 'know' the rules of their language, in the same sense everyone 'knows' their economic system. We all understand what money can do. But it is obviously ridiculous to take a mentalist approach. Nobody would call economics an organ of the mind.

Many sound patterns are the result of language being (like morality and economics) a self-organizing social institution, and are not generated at the level of individual behavior. We may start by considering those that result from the filling of a gap in a phonological system. It has often been observed that languages tend to fill holes in the patterns of their segmental inventories. Thus Antilla (1972) shows that Proto-Baltic Finnic had a system "in which the short vowels had one degree of height more than the long ones, and contained the only front rounded vowel in the whole system....(Modern) Finnish has filled every single gap and ended up with perfect symmetry." To take another example, it is not at all surprising for a language such as English, which at one stage had four voiceless fricatives /f, θ, s, ʃ/ and three voiced ones /v, ð z/, to acquire the missing voiced fricative [ʒ], as we have done recently. But it should be noted that this does not occur because of the biological drive acting on individual speakers and listeners. Filling holes in an abstract phonological system does not increase communicative efficiency for the individual. There is no increase in ease of articulation or auditory distinctiveness for any *existing* possible utterance. Nor, in general, does it help to identify the speaker in any way. (The exception is when the new sound is being brought in by borrowing from some other language or dialect. Using this sound may mark the speaker as wishing to identify with the speakers of the other language or dialect, as, for example, when a speaker of Southern British English starts using a voiceless fricative [x] in words of Scottish origin, such as 'loch'.)

This discussion of possible linguistic changes in terms of holes in phonological patterns is sometimes formulated in a slightly different way. There are a number of occasions when linguists talk about the segmental inventories of a language being such that they facilitate or hinder possible sound changes. Thus Maddieson (1984) suggests that if a language does not have /v/ or /β/ it has a greater possibility of developing a phonemically contrastive /β/ as a result of phonologizing an intervocalic [β] allophone of /b/.

Adding a new sound is like the emergence of a new species in biological evolution. It is possible to claim that it happens because God sees a gap and wants it to be filled so as to make the world more uniform. But an equally good claim is that if there is an ecological niche to be filled, events such as the random mutation of genes will conspire to fill it. In the same way a new phoneme is more likely to occur (to be borrowed or to be phonologized from an existing allophone), if it fits nicely into an existing pattern. This cannot be explained in terms of the behavior of individual speakers and listeners, just as the development of a new species is not due to the action of individual members of existing species. The communicative efficiency principle does not apply to how languages organise their sounds. When we discuss phonological systems we have moved to considering language not as part of an individual's behavior, but as a self organizing institution.

Many of the patterns currently described by linguists are patterns that occur simply in language considered as a social institution. In addition to the hole in the pattern phenomena, there are what Kisseberth (1970) calls phonological conspiracies. Dauer (1983) has assembled an excellent case for regarding stress timing in English in this way. Many people have observed an apparent tendency in English for stresses to recur at regular intervals of time. But it seems that this may be due to a fortuitous combination of circumstances. The fact that English words have a somewhat regular stress pattern, the possibility of alternative stress patterns in some words and of dropping stresses in some sequences of words, the reduction of weak syllables, and the clitic-like nature of many grammatical formations, all these things and more combine to lead to the appearance of stresses at appropriate intervals.

Perhaps the most startling conspiracy - one that seems to have deceived by far the majority of linguists - is the appearance of phonemes. Accounts of human behavior in terms of phonemes are nearly always examples of what has been called the psychologist's fallacy - the notion that because an act can be described in a given way that it is necessarily structured in that way. As will be shown in Chapter 4, phoneme size units play only a minor role in human behavioral acts such as normal speaking and listening.

Lindblom (1984) has suggested a nice analogy that can be extended to make this point clear. He has pointed out that termite nests appear to the outside observer to have a most intricate structure. There are great pillars and arches that rival those of medieval cathedrals. But it does not follow from this that individual termites know about arches. In fact they are simply following a very straightforward pattern of behavior, governed (in nest building) by a single rule: deposit grains of earth near other grains of earth that are scented with a termite secretion. At first this leads to random depositing of earth. But very soon the deposits are on top of other recent deposits and the pillars grow. As two pillars grow taller the scent is strongest on the sides closest to each other; and so those two sides grow together and form an arch. All from a single, simple, rule. Phonemes may be like arches in termite nests, visible to outside observers, but having no meaningful role in the activity of the individuals producing them.

Speech appears to be composed of sequences of segments because of the interactions of the different systems of which it is composed. The complex gestures involved in producing syllables have diverse parts that look as if they are categorically distinct. We call these diverse parts vowels and consonants, but we must always remember that these are just names for readily distinguishable aspects of the stream of speech. This point will be further elaborated in Chapter 4. Here we will simply note that those of us who have been exposed to an alphabetic tradition may be influenced so that we are very conscious of the possibility of describing speech in terms of vowels and consonants. But illiterates may have little or no concept of speech segments (Morais et al, 1979). Similarly those involved in adult literacy campaigns report that the concept of the segment is far from self evident. Intelligent adults who have been taught to write a few words cannot perform tasks such as naming other words that begin with the same segment. A language consultant who has been working extensively with a linguist will be able to learn the phonemic principle (Sapir 1949) just as a child can learn to read and write. But this is hardly evidence for phonemic units in the normal process of speaking and listening.

We can carry the termite analogy a step further still. Just because the individual termite cannot be considered responsible for the design of the arches and pillars in a termite nest, it does not follow that it is not interesting to describe these pillars and arches. They are a necessary part of the termite nest; when a pillar or an arch is needed to support the edifice which the community requires, if the nest is to be one which survives, then the requisite structure will be present. In that sense, a termite nest is self-organized. In a similar way a language gets the sounds that it needs to function as a viable communicative device. The segments and phonemes are present in the structure of that abstract entity, the language, considered as a social institution. They are units that are necessary for describing patterns that occur in this observable institution. Indeed, as Halle commented several years ago, 'Almost every insight gained by modern linguistics from Grimm's law to Jakobson's distinctive features depends crucially on the assumption that speech [or, in my terms, language] is a sequence of discrete entities.' (Halle, 1964.)

But despite the value of segments as descriptive units, it seems almost certain that the phonemic principle is not part of our genetic endownment (as it surely must be for those who view it as an innate ability). The manipulation of phonemes is an acquired ability. Evolutionists teach us that such things are properties of a culture, and not of an individual's physiology. The invention (not, for me, the discovery) of the alphabet occurred far too recently for it to have become part of our DNA. Indeed, as Gould (1981) put it: "*Homo sapiens* arose at least 50,000 years ago and we have not a shred of evidence for any genetic improvement since then......All that we have accomplished, for better or worse is a result of cultural evolution."

Our endeavors include building (like termites) social institutions such as language, morality, and economic systems. Each has, to a great extent, become its own thing, so that it is no longer entirely explicable in terms of outside forces. The evolution of language has involved its feeding upon itself, so that it must be described partly in terms of unique principles. The current accounts may not be correct. Descriptions of languages in terms of the fashionable metrical phonology (Halle and Vergnaud, 1980) are, after all, at least superficially very different from those of the older generative phonology (Chomsky and Halle, 1968). But the thread of the uniqueness of language that runs through them is still valid. And they are all, despite their authors' claims, descriptions of social institutions and not explanations of mental activities. Like termites who do not know how to build an arch, ordinary speakers and listeners do not know the sound pattern of English.

2. Why do we want to represent it?

2.1 The nature of phonological features.

The next question that we must consider more fully is what the phonetic representation is for. There are of course all sorts of possibilities. There are several different kinds of phonological statements one might want to make; and, as we have noted, phonetics being the scientific study of all aspects of speech, there are forms of phonetic representation that are not directly related to the phonology of the language being described. We will, however, leave these latter possibilites to a later discussion, and begin by examining the requirements that phonology places on a phonetic representation.

Phonology is concerned with explaining both the patterns of sounds that occur within languages and those that occur across languages. It must also provide a way of representing all the auditorily distinct lexical items within each language. In order to account for patterns of sounds it is necessary to arrange them into groups, by categorizing them in terms of phonological features. Phonological features therefore serve two different purposes: they group sounds together for the purposes of the rules that describe and elucidate phonological patterns, and they distinguish items in the lexicon. There is no intrinsic reason why the features required for the one purpose should be the same as those required for the other. In fact it is quite often said that certain languages (e.g. !Xóõ, according to Trail 1985) have little or no phonology — i.e. no alternations explicable in terms of rules — although they obviously have lexical items that have to be given distinct representations. Similarly, there maybe phonological alternations in a language that depend on some feature that is not actually used to distinguish words in that language. For example, in the Papuan language Yeletnye (Maddieson, personal communication) the feature Voice is not needed to distinguish lexical items. (Hawaiian and Maori are other more wellknown languages that are similar in this respect.) In Yeletnye there have to be rules that refer to the feature Voice in that the assignment of values of this feature in some positions depends on its values in other positions. For example, oral stops are voiced when surrounded by voiced elements, except stem initially.

These two aspects of phonological features are related to the two principal kinds of patterns that phonologies are required to explain: those that are concerned with the segmental inventories of languages, and those that are concerned with the types of rules that can occur. Many contemporary phonologists (e.g. Clements 1985, Halle 1988) are more interested in the types of rules that can occur than in phonological inventories. But it seems to me that an equally important task for phonological theory is to explain why languages have certain contrasting sounds and not others. One of the major pieces of work of this kind published recently is that of Maddieson (1984). A proper theory of phonology must account for both the inventories that Maddieson observes and the rules that other phonologists find in the world's languages.

In order to construct such a theory we must consider how languages get to be the way they are. Languages are the products of speakers and listeners; and the acts of speaking and listeners leave their mark on languages in different ways. One of a speaker's goals is to communicate without undue articulatory effort. A listener requires a sufficiently distinct sequence of sounds to be able to get the message in a sufficiently short length of time. Within these conflicting aims there are a number of compensations possible between articulatory effort, auditory distinctiveness, and rate of speech. Usually the speaker is able to take the initiative in setting this balance. Only occasionally does the listener have to interrupt and ask for clarification in some way. The balance between the conflicting forces is clearly different in different circumstances. A speaker addressing an unfamiliar audience on a complex topic may talk slowly with a careful articulation. Two close friends exchanging information will be able to talk more quickly and with a considerable reduction in the degree of precision of the articulatory movements. On some occasions, when the listener almost certainly knows the words to expect, virtually no distinctive articulations may be needed. Lovers in bed need few precisely spoken words; and soldiers on the parade ground can interpret the wordless bellowing of the drill sergeant with great rapidity.

In general, as listeners become more familiar with particular words or phrases, speakers will be able to use more articulatory assimilations. This topic has been well enough covered in the vast literature on historical phonology to need little further documentation here. We might, however, note that as well as obvious assimilations that occur in pronouncing items such as 'handkerchief' and 'in between' as ['hæŋkətʃif] and [ɩm bə'twin], many cases of apparent *dis*-similation are actually examples of economy of effort. This is the case for Grassman's law, which states that the first of two aspirated stops in a word will become deaspirated (so that, for example Indo-European *thrikhos becomes classical Greek trikhos 'hair'). This can be interpreted as dissimilation, the consonants in a word become more unlike one another. But it is also an example of economy of effort. Aspirated consonants are very distinct from all other sounds (Singh and Black, 1966); but they are also costly in that they use considerable respiratory energy. A word with two such sounds is very costly, and an obvious candidate for pruning in any attempt to reduce the overall effort required for an utterance. Ohala (1981) has given a convincing account of dissimilation in terms of the listener as the source of sound change. I would only add to his account that in Grassman's law the conditions were ripe for it to occur because it satisfies the biological drive for economy of communication (not that Ohala would put it that way).

The notions of auditory distinctiveness and economy of articulatory effort are both important in the formation of phonological patterns. Because sounds that were produced in similar ways usually have a similar acoustic structure, features that group sounds in articulatory terms will usually also group them in acoustic terms. But the reverse is not always true; some segments can sound very similar although they were produced in quite different ways. There are thus two different kinds of phonological features: those that have an auditory basis, and those that group sounds in articulatory terms.

We must also note that some phonological patterns are the result of historical processes which were, when they occurred, due to articulatory or auditory aspects of the sounds in question, but which now group sounds in ways that cannot be explained in terms of how they are currently heard or produced. Consequently the patterns that result from them may be associated with *ad hoc* features within languages that have neither an auditory/acoustic nor a physiological basis. An example of this kind of phonological description occurs whenever the vowels of English are described in terms of the feature Tense. This is a notoriously difficult feature to define; but it nevertheless specifies a very real mental grouping that has to be considered part of the sound pattern of English. We can substantiate the psychological reality of features such as Tense (at least for literate speakers of English) by showing that they are needed in productive rules, or by phonological experiments as described by Ohala and Jaeger (1986). Within the description of a particular language, *ad hoc* features should be used only when they are psychologically real.

Another kind of phonology might prohibit the use of a feature that had no physical basis on the grounds that a phonology of this kind is too unconstrained, and too difficult to learn. Given the possibility of *ad hoc* features of the sort described above it would be possible to devise abstract phonologies that described all sorts of pretty patterns that might have no physical correlate of any kind. It is arguable that such phonologies might be hard or even impossible to learn. If there are no physical cues for groupings of sounds, how could children, who have no access to the history of the language, internalize patterns that have only a historical basis? The answer is that from the age of 6 or 7 on some English-speaking children at least do have access to the history of the language. It is represented in the orthography. There is considerable evidence (Moskowitz 1973, Ohala 1974, Jaeger 1986) that people can use orthographic knowledge as the basis for forming phonological classes. The end result is that the phonological features become mental objects defined by the phonology itself. They function, like the square root of -1 in mathematical formulae, as useful concepts that have no physical reality but have great value in explaining the way things work.

Any non-physical arbitrary groupings of segments that simply reflect historical events in a particular language or group of languages are, by definition, not part of a universal phonetic feature set. When we are concerned with the general nature of human language, then we must consider only phonological descriptions made in terms of phonetic features that have a physical basis that could apply to all languages. Comparative phonologies that are restricted in this way are, of necessity, different from those of a single language in which *ad hoc* features might be permitted.

2.2 The need for auditory/acoustic features

As we have noted, the physical basis for phonological features may be physiological or acoustic. There are several important natural classes that are the result of sounds having an acoustic structure such that they have certain auditory properties in common. It is somewhat ironic that this great insight of the Prague school, much touted by Jakobson, Fant and Halle (1951), should now be overlooked by the phonologists who are their successors. The present situation arises partly because of the view of phonology in SPE (Chomsky and Halle 1968), in which features are considered to be mental entities. From this point of view it is just a matter of exposition as to whether features are defined in articulatory or acoustic terms. But this is simply not true. Segments get grouped together into natural classes not because of some general mental property, but because of specific properties relating to the way sounds are heard, or to the way they are produced. Of course all features have both articulatory and acoustic properties in the somewhat irrelevant sense that as soon as a feature has been defined in one way or the other it can be regarded as a linguistic unit that characterizes the lexical items of a language. These lexical items have to be capable of being both spoken and heard; and the features that characterize them must have both kinds of properties. But it does not follow from this that we should consider the linguistic function of a feature as being *required* in both domains or that we can define it equally well in either way. (A similar view has been expressed by Lieberman (1970), though he also claimed that languages prefer features defined in both domains.)

I do not want to overstate my case in this matter; and, accordingly, in Table 2.1, which lists some auditory features, I have put at the the top of the list a feature, Voice, which has both kinds of characteristics. For many aspects of sounds there are salient auditory and physiological properties, either one of which could be regarded as the basis for an observed grouping. In each of these cases we should we define the feature in both auditory and articulatory terms. Thus the feature Voice can be defined as having regular vibrations of the vocal cords and as having an acoustic structure with well defined harmonics. But, as is exemplified in the lower part of Table 2.1, there are other features that have no phonologically useful articulatory correlates.

Table 2.1 An example of a feature Voice, that can be defined in both auditory and acoustic terms, and a set of features that have non-specific articulatory properties, but determine auditorily based natural closures.

Feature	Traditional terms	Possible values	Brief description
Voice	Voiced	[+voice]	periodic low frequency energy /vibrating vocal cords
	Voiceless	[-voice]	absence of such energy /not vibrating vocal cords
Grave		[+grave]	aperiodic low frequency energy
		[-grave]	absence of such energy
Sibilant	Sibilant	[+sibilant]	aperiodic high frequency energy
		[-sibilant]	absence of such energy
Height	High vowel	[high]	low F1
	Mid-high	[mid high]	mid low F1
	Mid-low	[mid low]	mid high F1
	Low vowel	[low]	high F1
Brightness		[+brightness]	high (F2'-F1)
		[-brightness]	low (F2'-F1)
Sonorant		[+sonorant]	periodic energy with well-defined formants
		[-sonorant]	absence of such energy

We will now consider the auditorily based features shown in the lower part of Table 2.1. The main point in this chapter is just to show that there are these two kinds of features. As nobody seems to doubt that some, if not all, features can have an articulatory basis, the discussion of these features will be left to a later chapter. One of the most well known auditory features is Grave, which groups some labial and velar sounds in accordance with their spectral characteristics. Sounds such as [p,k,f,x] are produced in very different ways, but they sound similar because they have a comparatively large amount of aperiodic acoustic energy in the lower part of the spectrum. This similarity is reflected in historical changes such as English [x] to [f] in words such as 'rough, tough,' or the parallel phenomena in Danish whereby "lugt" becomes "luft" (Basbøl 1974). These changes are completely inexplicable in articulatory terms. Phonological alternations involving the feature Grave have been described by Hyman (1975) who shows that different vowel allophones occur before Grave consonants in Feʔfeʔ. Other cases of phonological alternations include spirantization in Hebrew triggered by the Grave consonants /p, b, k/. There seems to be no doubt that Grave consonants form a natural class.

Chomsky and Halle discarded the feature Grave because they found it did not provide a satisfactory basis for characterizing differences in place of articulation. This is undoubtedly true; from an articulatory point of view the feature Grave does not distinguish the appropriate natural classes. But this does not mean that it fails to characterize a natural class of sounds from an auditory point of view. Throwing out Grave just because it does not have a useful articulatory correlate is as bad as it would be to throw out Nasal just because it does not have a unified set of acoustic correlates that form a basis for a natural class.

Note that the feature Grave as defined in Table 2.1 is not exactly the same as the feature proposed by Jakobson, Fant and Halle 1951). Their definition was "the predominance of one side of the significant part of the spectrum over the other." It was intended to include both consonants and vowels. The feature Grave as defined here is in practice restricted to obstruents (and, perhaps, voiceless approximants) because it stipulates that the auditory characteristic of a Grave sound is that there is salient *aperiodic* energy in the lower part of the spectrum. In speech, this type of energy occurs only in stop bursts and fricatives (and, perhaps, a voiceless labial-velar approximant). There is no auditory property of this sort that links particular vowels with particular consonants. (But, as we will see later, there are links between particular vowels and consonants specified by certain articulatory features.)

Note also that this definition of Grave implies that [- grave] sounds are not necessarily Acute in the old Jakobsonian sense. All sounds that do not have a significant amount of aperiodic energy in the lower part of the spectrum are [- grave], irrespective of whether they have a significant amount of aperiodic energy in the upper part of the spectrum or whether they do not have any aperiodic energy at all.

Another auditory feature that is of importance in grouping consonants I have here called Sibilant, following the traditional phonetic usage. It is not exactly equivalent to the Jakobsonian feature Strident in that the feature Strident has also been used to distinguish [f,v] from [ɸ,β], thus resulting in the rather unnatural class of strident sounds [f,v,s,z,ʃ,ʒ]. So as to make make the difference in definition plain, I have retained the traditional term Sibilant, which has long been used (e.g. by Holder 1669, and many phoneticians after him) to identify the class of sounds [s,z,ʃ,(ʒ)].

It is interesting to consider whether it might be possible to give an articulatory definition of this feature, in that Sibilant sounds are always pronounced with the jaw raised so that there is a very narrow gap between the upper and lower front teeth. The high frequency aperiodic acoustic energy that gives rise to the auditory characteristics of this feature is due to the jet of air striking this narrow gap (Catford 1977a, Shadle 1985). However, the fact that sibilant sounds have an articulatory attribute in common is an unlikely cause for their acting together in historical changes and morphological alternations. There is no evidence showing that jaw position is a salient characteristic of sounds causing them to be grouped together, whereas the auditory grouping of these sounds is evident in the perceptual confusion data of Miller and Nicely (1955) and its reanalysis by Shepard (1972), and in the perceptual similarity judgments reported by Ingram (1975).

It is appropriate at this point to consider further what is at issue in claiming that a certain feature (e.g. Sibilant) should be defined in auditory rather than acoustic terms. It is not a matter of whether there is or is not a feature of this kind. There is little doubt that sibilants form a natural class of sounds that act together in phonological rules. Nor is it a matter of formal evaluation of rules. Given that there is a feature Sibilant the system for evaluating its use within a phonology will be the same irrespective of its phonetic attributes. What is at stake is whether the auditory definition provides a better explanation for the grouping than a definition in terms of the articulatory attributes. Until there is some evidence for the shared articulatory properties being the reason for this grouping, it seems preferable to continue to maintain that the well attested salient auditory characteristics are the basis for the natural class.

The most outstanding features of the auditory type are properties of vowels. A problem that arises in discussing these features is that it has not been generally recognized that vowels have both articulatory *and* auditory properties. What may (loosely speaking) be called the height and backness of the body of the tongue are important considerations in accounting for articulatory interactions between consonants and vowels. Descriptions of these interactions will require us to use features that actually specify articulations. The two articulatory features needed for these purposes are similar to, but not the same as, the auditory features that account for some of the auditory interactions among vowels. Unfortunately the same names have been used for both the articulatory and the auditory features. I will use the terms High — Low, and Front — Back to refer to articulatory properties. I will use the terms Height and Brightness to refer to the auditory properties of vowels.

We are not used to thinking of vowels simply as sounds to be judged in terms of auditory qualities. Our difficulties are further compounded by the fact that the auditory features we have considered so far, such as Grave and Strident, operate in a binary way. They are either present or absent, and we do not have to think in terms of scales, as we do for vowels. We all realize that the notes on a piano are arranged on a scale going from low to high. But most people find difficulty in thinking of any properties other than musical scales on which sounds can be ordered; loudness is the only other auditory property that is generally recognized. Nevertheless it is a fact that some pairs of sounds - [i] and [e] for example - are more alike than others - for example [i] and [a]. Accordingly there must be some property or properties of vowel sounds that enable listeners to make these judgements. These properties may, of course, be in the articulatory domain. It may be that listeners are simply referring their auditory impressions to their tacit knowledge of the articulations required to produce vowels, as suggested by the motor theory of speech perception (Liberman and Mattingley, 1985). But before we assume that this is the case, we should consider the auditory judgments that people make, and see whether they are in fact equivalent to their articulatory gestures.

A necessary first step in describing how the vowels [i e a o u] differ from one another in auditory terms is to decide how many auditory properties are involved. It turns out that from an auditory point of view, the five vowels [i e a o u] differ in terms of two auditory properties. If we ask observers to listen to pairs of vowels from this set, and to say which pair of vowels is furthest apart, which is most alike, and so on, we can obtain measures of the auditory distances between each of them. Terbeek (1977), who performed an experiment of this sort on a larger set of vowels, found that the sum of the auditory distances between [i - a] and [a-u] is greater than the distance between [i-u], indicating that there is more than a single auditory property with [i-a-u] being ordered along it on a straight line. This is not a surprising finding for most phoneticians. But what is a much more interesting question is whether the basic vowel system [i e a o u] varies in more than two distinct auditory properties. It may well be that judgments of auditory similarity of just the vowels [i e a o u] can be represented perfectly adequately without having to take three independent dimensions into account.

A possible pair of auditory properties maybe Height and Brightness. Height is a well known term for describing an aspect of vowel quality, but again I would like to emphasize that here it is being regarded as simply an auditory property. As such, it cannot be defined in anything other than impressionistic terms reflecting a listener's judgments. It is like pitch, for which the only definition offered by the Acoustical Society of America is that it is that property of a sound that enables it to be ordered on a scale going from low to high. In the case of vowel Height the definition we would offer is that it is the property that enables a vowel to be ordered on an auditory scale of vowel height going from low to high. But just as pitch can be related to fundamental frequency, so can vowel Height be related to measurable acoustic parameters. It is generally agreed that the acoustic features that listeners attend to are formant frequencies and amplitudes. The major acoustic correlate of the auditory property vowel Height is the frequency of the first formant (F1).

The other auditory property has been termed Brightness ("Helligkeit") following Trubetzkoy (1929, 1939), and more recently Fischer-Jørgensen (1985). The acoustic correlates of Brightness are not fully established (it is another topic on which more research remains to be done), but we will take it to be (F2'-F1), the difference in frequency between the first formant and F2', a form of the second formant frequency modified so as to account for the influence of higher formants. Suitable formulae for calculating F2' are discussed by Bladon and Fant (1978).

The terms Height and Brightness are related to the traditional articulatory terms in a complex way. Brightness is a combination of all three articulatory vowel features, High — Low, Front — Back, and Round; it is not just a matter of conflating backness and rounding. High front unrounded vowels have the highest value of Brightness, low back neutral vowels have a mid value and high back rounded vowels have the lowest value.

The explanatory power of the two auditory features for vowels can be exemplified by reference to notions of vowel raising, as will be discussed in the next section. Here we will consider a different point, namely, the way in which the auditory aspects of vowels can explain the dominance of the five vowel system [i e a o u]. From an auditory point of view, the fact that these particular vowels are selected is not surprising. As suggested by Ladefoged (1971:76) there is probably "some kind of principle of maximum distinctiveness whereby the auditory differences between the vowels in a language tend to be kept at a maximum." Liljencrants and Lindblom (1972) showed that if the auditory properties in question depend on the first two formants, then the most likely 5 vowel set would be [i e a o u]. Further work (summarized by Lindblom 1986) confirms this view.

Languages as diverse as Swahili, Spanish, and Hawaiian have five vowels, with qualities something like [i e a o u]. These and only these vowels are used by approximately 20% of the world's languages (Maddieson 1984). On the basis of the material reviewed in the last chapter, we can conservatively estimate that there are at least 50 distinct vowels (neglecting differences in phonation type, etc) that languages might have chosen as the vocalic elements of their segmental inventories. The likelihood of the same five being chosen so frequently is therefore comparable with the likelihood of playing poker and finding that one hand in five always had the Ace, King, Queen, Jack, and Ten of Spades. It is therefore an absolutely astounding fact that so many language have the vowels /i,e,a,o,u/, and any theory of phonology that does not offer an explanation for this fact must be considered to be seriously lacking. We can now see a two part explanation for this incredible fact. The first part suggests that the optimal number of vowels that a language needs to have is in the neighborhood of five. Three vowels (or even less) might be sufficient for conveying an appropriate number of distinctions among possible lexical items, but we may hypothesize that this is not optimum, and puts some strain on the system. More than five vowels, on this hypothesis, may be more than needed for optimal communicative efficiency. Given that five vowels is the modal number then the second part of the proposal is that the distinctiveness principle applies to the auditory features, making the most favored set [i e a o u].

From an articulatory point of view, there is no reason why front unrounded and back rounded vowels should be more common than the reverse combinations. From an auditory point of view, the choice of these particular vowels is part of the tendency among languages to maximize perceptual contrasts. Phonologists who regard all features as having only articulatory definitions have no explanation for the remarkable facts of vowel distribution. But it should be emphasized that this does not mean that articulatory features do not have a role in the description of vowels. The action of the body of the tongue in the production of a vowel is specifiable in terms of physiological features that are also applicable to consonants (and thus show the relations between vowels and consonants). There should be no doubt that in order to form the correct phonological classes of vowels these sounds have to be characterized in *both* physiological and auditory terms.

The next auditory feature listed in Table 2.1 is Sonorant, a feature which I define in a slightly different way from the definition in SPE. In many languages sounds such as [m, n, l, r] act together as a class. For instance, in English these sounds are syllabic after a stop or a fricative, as at the ends of the words 'table, tassle, sudden, prism, hidden', but not (for most of us) after after other sonorants as in 'film, kiln'. The feature Sonorant is hard to define meaningfully in articulatory terms. The notion 'spontaneous voicing' (Chomsky and Halle, 1968) does not get at the essence of what it is that causes vowels, nasals, laterals and some approximants to be grouped together. Better articulatory statements can be made in terms of the function of the articulatory system as a whole: sonorant sounds are those in which the vocal cords are vibrating and there is no significant build up of oral pressure. But if we are to claim that the feature Sonorant has this kind of articulatory basis, then we must claim that vocal cord vibrations and lack of pressure are both sensed by a speaker, and then combined so that together they are considered to form a salient psychological percept. This is a rather far-fetched notion for which there is no evidence. The fact that a feature can be defined in a certain way does not necessarily mean that this definition is any help in explaining *why* the feature groups sounds together into a natural class. Sonorant sounds are clearly related by having a periodic, well-defined, formant structure. They behave the same way within a language not because they are made alike, but because they sound alike.

Similar observations can be made about the feature Rhotacized, which is associated with a lowering of the frequencies of the third and fourth formants. As has been shown by Lindau (1985) many forms of *r* share this auditory characteristic, although they may have been produced by very different articulatory means. There is, however, a complicating factor in that Lindau also shows that some forms of *r* do not have this acoustic structure; they are linked to some (but not all) of the other *r* sounds by having a similar articulation. This observation leads us to another very important consideration concerning the nature of phonological features which we will examine in the next section.

2.3 Multi-propertied terms

So far we have suggested that there are three kinds of phonological features: those that reflect some *ad hoc* grouping of sounds resulting from diverse historical processes; those that can be associated with a physiological property (which will be discussed more fully in the next chapter); and those that can be associated with an auditory property. When we are concerned with general properties of human language, then we can use only phonological features that reflect a physical property of one or other kind; and that property must be the same in all languages. There is a long and valid tradition to this effect in feature theory. Jakobson, Fant, and Halle (1951) state: "A distinctive feature cannot be identified without recourse to its specific property." Chomsky and Halle (1968) make it quite plain that "... each feature is a physical scale defined by two points." More recently this view has been reaffirmed by Stevens (1983): "Each feature is represented in the sound wave as a unique acoustic property to which the auditory system responds in a distinctive way." Note that this view is even stronger than Chomsky and Halle's, in that it claims that there is a unique *acoustic* correlate of each feature. But for all these authors, each segment, S, is composed of a number of features, F, each of which is related to some physical parameter, P, as sketched in Figure 2.1.

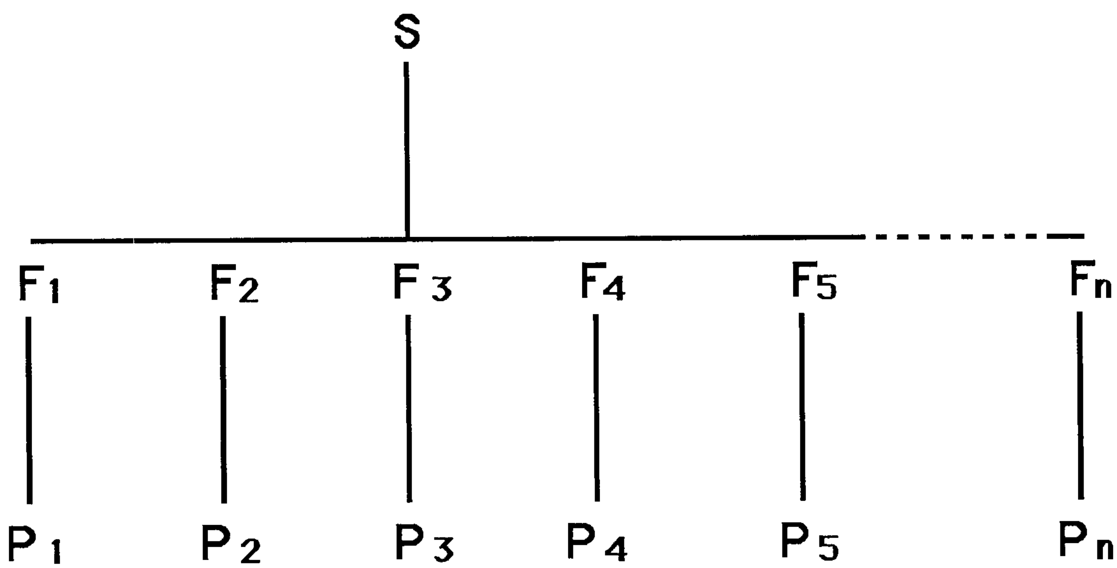

Figure 2.1. The relationship between a segment (S), features (F), and physical parameters (P) in traditional feature theories.

There are several reasons why it seems advisable for a general linguistic theory to permit only a single correlate for each feature. The first is that, as we have noted, without this constraint, features do not have any simple universality. Thus, unless we know that a particular feature always involves a particular attribute, we cannot consistently equate descriptions of different languages involving that feature; and if we cannot equate descriptions we cannot say which phonological phenomena are linguistic universals. Another reason for requiring a fixed attribute for each feature would be that it is in accordance with a view of human perception that holds that there are specific feature detectors, as proposed by Stevens and Blumstein (1975), Blumstein and Stevens (1975), and Stevens (1983).

It is worth digressing slightly to discuss this view more fully. If it were correct, it would indeed be a compelling reason for regarding a feature as reflecting a single acoustic property. However, other than in the cases of the auditory features such as those discussed in the preceding section, there is no case reported of a unique acoustic property that is an invariant correlate of a phonological feature. According to Stevens and Blumstein (1981): "The theory of acoustic invariance has been elaborated most completely for place of articulation in stop consonants." But, without disparaging Stevens and Blumstein's considerable achievements, their success rate is not very high. They suggest that there is a particular relationship between the spectral burst and the formant frequencies for each place of articulation. But they find the properties proposed for detection of place of articulation are present in only about 83 % of initial stops, 75 % of final stops if exploded, and 77% of initial nasal consonants. Further work by Lahiri, Gewirth, and Blumstein (1984) led to some improvement in these percentages. By allowing the invariant acoustic properties to be both time varying and relative they were able to distinguish bilabial and dental/alveolar stops more than 91% of the time. But this improvement in distinguishing places of articulation is applicable to only one manner of articulation (stop), and only one place in the syllable (initial). Phonological features group sounds in terms of place irrespective of whether they are stops, nasals, or fricatives, and irrespective of whether they are initial or final in the syllable. If correct detection of place of articulation could be achieved in 90% or more of both initial and final positions for stops, fricatives, and nasals, then the feature-detector hypothesis could be considered more seriously. There have been suggestions (Stevens and Blumstein 1981, Stevens 1983) of possible invariant properties for several other phonetic categories. However it is completely clear that there are no reports of the analysis of natural speech that show the presence of invariant acoustic properties, however complexly defined, for a wide range of phonological features.

It is always possible to argue that the failure to find an invariant acoustic property for a particular feature is due to our lack of diligence in looking for it. The claim could be that the invariance is there, but we just have not found it. This is rather like saying that part of the moon is made of green cheese. This is a very difficult claim to disprove. When we all fail to find the green cheese, it could be said that this is just because we have not looked in the right part of the moon. Of course, there are many occasions in scientific endeavours when it is appropriate to seek structures that are almost certainly present but very difficult to find. This seems to be the situation at the moment in genetics where people are searching for the correlates of inherited characteristics in the structure of the DNA. But the search for invariant acoustic properties of all phonological features is not like that.

There is no reason to expect all phonological features to be in a one-to-one relation to acoustic scales. It is much more likely that the articulatorily based features will be associated with variations in many acoustic parameters. This view of phonological features is illustrated in Figure 2.2. A segment is still described in terms of features, where the features represent convenient labels for classes of sounds needed by the phonology. Some features (those that reflect auditorily based natural classes) have simple, direct acoustic correlates; but some features are connected to acoustic parameters in a complex, many-to-many relationship. Thus the auditory features 1 and 2 are each associated with only a single property, as is feature 3; but the articulatory feature 4 is also somewhat associated with acoustic property 3, as well as having its primary association with acoustic property 4, and some association with acoustic property 5; and so on.

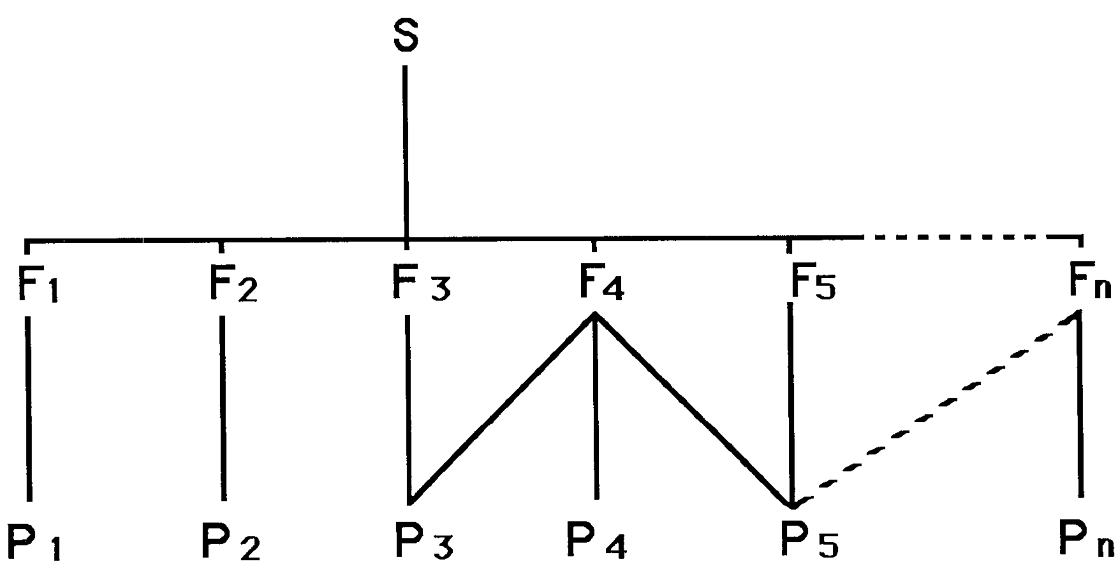

Figure 2.2. The relationship between a segment (S), features (F), and physical parameters (P) proposed here.

If there were always a many-to-many mapping between phonological features and both articulatory and acoustic parameters, the phonological features themselves would always be abstract properties, and thus useless for making generalizations about the nature of language. They would all be like the features based on historical groupings of sounds which we discussed earlier, in that they would be entities that are completely defined only within a particular phonological description. In the system we are considering, each feature must have one property, which may be acoustic or articulatory, that serves as a defining characteristic.

There are occasions when this may limit what we want to call a single type of sound in what seems an undesirable way. Consider, for example, the class of *r* sounds which we were discussing immediately prior to this section. Each member of this class of sounds resembles another member with respect to some property, but it is not the same property that constitutes the resemblance for all members of a class. Trills and taps resemble each other in the duration of the pulses involved, tongue-tip trills and uvular trills are similar in the patterns of consecutive pulses, and tongue-tip taps and tongue-tip approximants are close to each other on an articulatory stricture scale. It is thus possible to relate uvular trills to tongue-tip approximants, provided one proceeds in steps across the members of the class of r-sounds. (The situation is even more complex in the acoustic domain, where it is not possible to find even this kind of resemblance. Only the tongue-tip sounds tend to have a lowered F3, and even they may not, as in the Spanish sounds analyzed by Delattre (1965), and Degema taps analyzed by Lindau (1985). Uvular rhotics will have a raised F3 approaching F4.)

The relationship between sounds of the *r* type and their phonetic properties is more of a "family resemblance" kind (Wittgenstein 1958) than of a kind that can be defined in terms of phonological features. In such groupings we find that family member S_1 resembles S_2, which resembles S_3, which resembles S_4, etc. Although family members S_1 and S_4 may not share any essential phonetic properties, it is possible to express their relatedness as a set of steps across other members. Different sounds in different languages can be described as all being of a certain type; but in the view being developed here they are characterizable in terms of a single phonological feature only if there is a single physical property that constitutes the phonetic correlate of that feature.

Lindau and Ladefoged (1986) have given other examples of sounds which are of the same type but which I would now consider to be not definable in terms of a single phonological feature. One of their examples is the family of sounds that is involved in vowel harmony in African languages and the register distinction in Mon Khmer languages. Sets of sounds distinguished by vowel harmony differences and sets distinguished by register differences have a number of characteristics in common. The distinctions often involve several articulatory parameters which are related to one another, such as tongue-root movements, tongue height variations, vertical larynx movements, and changes in laryngeal tension (and hence in phonation type). But these factors are not all describable in terms of a single feature.

Across languages there are often situations that we regard as partially the same, in the family resemblance sense. We undoubtedly need some way of expressing these similarities. It seems that this can best be done by grouping features together so as to show parallels in the phonological processes that occur in different languages. As will be shown in the next chapter, we can do this by arranging the features in a hierarchy in which the higher nodes, rather than the features themselves, may be used to refer to more general phonological processes. In this way we can avoid weakening the notion that each terminal feature must be definable in terms of a single scale, in either the articulatory or the acoustic domain. There may be cover terms or higher nodes that have multiple properties, but terminal features have only a single phonetic correlate.

2.4 Multi-valued features.

There is one further formal consideration affecting both articulatory and auditory features that must be clarified before we go on, in the next chapter, to complete the set of phonological features by describing the required articulatory features. This is the question of whether all features have to be binary. Implicit in my proposal for the auditory features Height and Brightness is the notion that these features are non-binary; each of them permits a range of values along a scale. Thus the feature Height has (at least) three possibilities: [high], [mid] and [low]. The process of going from [low] to [mid] is the same as the process of going from [mid] to [high]. Only with this notion of scalar features can we maintain the notion of a two-dimensional (as far as these features are concerned) vowel space. Without this notion there is little likelihood of a true explanation not only of why vowel systems are as they are, but also of many phonological processes such as the English vowel shift and the changes in vowel quality that are due to stress reduction. It is, of course, possible to describe all such changes in vowel quality in terms of binary features as illustrated by Chomsky and Halle (1968). But any rule using binary features that has to account for [+high] becoming [-high] in the same circumstances as [-low] becomes [+low] inevitably misses a linguistically significant generalization. Simply by virtue of having to use two separate features it cannot show the unity of the process that is expressed by a rule of the form: [n high] → [n+1 high]. Why should the counterpart of : [+high] → [-high] be [-low] → [+low] rather than say, [+low] → [-low], or [-back] → [+back]? Similarly, if High and Low were truly independent binary features it would be difficult to give an explanatory account of vowel reduction in which both [+high] and [+low] vowels become simultaneously [-high, -low]. Current feature theories achieve the correct result by defining High and Low as deviations in opposite directions from the same starting point, a mid vowel. They also have to have a marking convention that prohibits a vowel from being simultaneously [+high] and [+low]. All this is equivalent to arranging, in a rather cumbersome notation, that [high], [mid] and [low] form an ordered set of values on a single scale. So why not say just that?

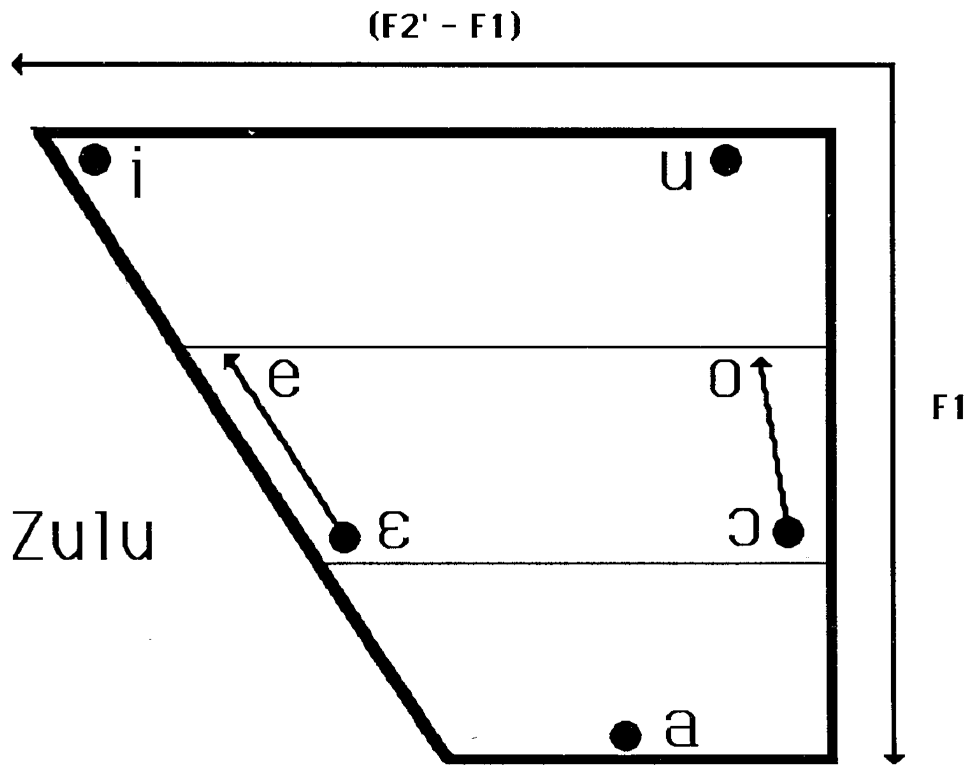

Figure 2.3. Vowel raising in Zulu.

Good examples of vowel raising can be found in a number of languages spoken in Southern Africa. The phonetic facts are hard to state in terms of the conventional SPE features, which permit the specification of only three vowel heights. In the Nguni languages there are five vowels, which, in traditional IPA terms, have the qualities [i,ɛ,a,ɔ,u]. Each of these vowels is fairly similar to the corresponding cardinal vowel, except [a], which is retracted so that it is in between the cardinal [a] and [ɑ]. Preliminary data indicate that the unmodified qualities of these vowels are as shown by the solid points in figure 2.3 The two mid vowels [ɛ,ɔ,] have raised variants that occur whenever the following vowel in the word is [i] or [u]; their phonetic qualities are as indicated by the arrow heads in the figure. (The phonological conditions can be further elaborated, but this is sufficient for our purpose here.) We therefore have to have a rule of the form:

[mid-low] → [mid-high] / (C) [high]

This rule could be formulated using the SPE features High, Low and Tense, but this would be a purely arbitrary use of Tense to have exactly the same phonetic exponents as the features High and Low; the vowels [e,o] differ from [ɛ,ɔ] in exactly the same way, physiologically and acoustically, as [ɛ,ɔ] differ from [a].

The situation is complicated still further in languages of the Sotho group, which have seven underlying vowels [i,e,ɛ,a,ɔ,u]. In these languages all four of the mid vowels [e,ɛ,ɔ,o] have raised variants in similar circumstances to the raised variants in the Nguni languages. Preliminary data indicate that the phonetic facts are as summarized in figure 2.4. It seems obvious that we need a multi-valued feature vowel Height to make the correct generalizations.

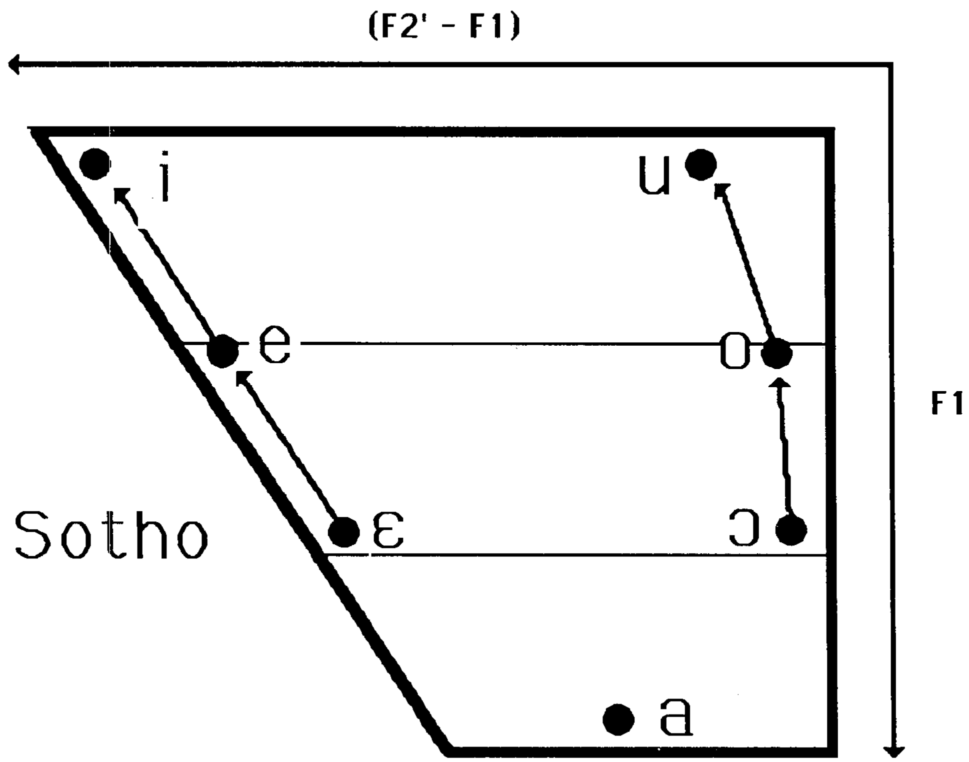

Figure 2.4. Vowel raising in Sotho.

The whole question of vowel specification is further complicated by the fact that what I have characterized as the difference between a binary and a multi-valued approach, others (e.g. van der Hulst 1988) regard as being binary versus single-valued. For me, multi-valued means that a single phonological feature has the possibility of one, and only one, value out of an ordered set of multiple values, ranging from a minimum amount of the property in question to a maximum. For those who use the notion of single-valued vowel features, each of these features is regarded not as a variable that can have different values, but as a component whose presence indicates a directional pull. It then becomes possible to denote the strength of the pull in a given direction by adding further components. The notation [i-a] (or some equivalent way of showing dominance) would indicate a vowel midway between [i] and [a]; but [ii-a] would indicate a vowel nearer to [i], and [i-aa] one nearer to [a]. This certainly makes evident the two dimensional nature of the basic vowel space; but it is not clear that it can capture the correct phonetic generalizations in the vowel raising cases discussed above.

Further examples will be provided in the next chapter of multi-valued articulatory features in which there are at least three mutually exclusive values, arranged in an ordered series. It seems to me that the evidence that some phonological features are sometimes multivalued is simply overwhelming. Many are binary; but in the case of some features we need scalar values both in phonological rules and in accounting for universal patterns of sounds among languages. There is no basis for the argument that all features must be binary because phonological rules have been shown to operate in terms of such features. Phonological rules have never been entirely binary; about one third of the rules in SPE — those involving Stress — use a multi-valued notation that permits scalar quantities. The fact that much of the patterns of sounds within languages can be described in binary terms does not show that it is not equally possible to use another formalism.

I hope that phonologists will soon recognize that we are due for our own little Copernican revolution. It is possible to describe many of the observations of astronomy, and to predict future eclipses while still maintaining that the earth does not move and the sun goes round it. But as Galileo (1633) whispered after being forced to retract his Copernican heresies "Epur si muove [still it does move]." Like Galileo, I will not go to the stake for my belief. Still five vowel systems are most favored. Mid vowels are between high and low vowels. An interface between phonetics and phonology must allow some phonological features to have non-binary values.

Hyper feature	Major node	Terminal Feature	[named term]	Traditional term
			[protruded]	*Rounded*
	Labial	Labial	[compressed]	*Neutral*
			[retracted]	*Labiodental*
			[laminal]	*Laminal*
		Apicality	[apical]	*Apical*
	Coronal		[sublaminal]	*Retroflex*
			[dental]	*Dental*
Place		Anteriority	[alveolar]	*Alveolar*
			[postalveolar]	*Postalveolar*
			[front]	*Front* *Palatal*
		Back	[central]	*Central*
	Dorsal		[back]	*Back*
			[high]	*High vowel / Velar*
		High	[mid]	*Mid vowel / Uvular*
			[low]	*Low vowel / Pharyngeal*
	Radical	ATR	[+ ATR]	*Advanced Tongue Root*
			[- ATR]	*Epiglottal*
			[stop]	*Stop*
		Aperture	[fricative]	*Fricative*
Stricture			[approximant]	*Approximant*
		Trill	[+ trill]	*Trill*
			[- trill]	
		Lateral	[+ lateral]	*Lateral*
			[- lateral]	*Central*
Nasality		Nasal	[+ nasal]	*Nasal*
			[- nasal]	*Oral*
		Voice	[+ voice]	*Voiced*
			[- voice]	*Voiceless*
			[closed]	*Glottal stop*
			[creaky]	*Creaky / Laryngealized*
		Glottal aperture	[modal]	*(Modal) Voice*
Laryngeal			[breathy]	*Breathy / Murmur*
			[spread]	*Voiceless*
		Aspiration	[+ aspiration]	*Aspirated*
			[- aspiration]	*Unaspirated*
		Pitch -	(tone and intonational features - not considered here)	
	Pulmonic	Pulmonic	[+fortis]	*Fortis*
			[- fortis]	*Lenis*
Airstream -	Glottalic	Glottalic	[implosive]	*Implosive*
			[ejective]	*Ejective*
	Velaric	Velaric	[+ click]	*Click*
			[- click]	

Figure 3.1 A hierarchical arrangement of phonological features.

3.1 The hierarchy of articulatory features

Patterns of sounds occur in languages largely because of the way in which sounds are made. There is no doubt that the principal way of grouping sounds together in a phonological description is in terms of their articulatory correlates simply because the articulations necessarily determine the acoustics (although, as we have seen, the other way round is not true). Figure 3.1 presents a summary of the articulatory features, arranged in a hierachical structure.

The notion of a hierarchical structure for phonological features has been firmly entrenched in phonological theories, at least since the seminal paper by Clements (1985). This notion has been used to account for constraints on phonological processes such as the spreading of a feature from one segment to another. But it seems to me that the concept of a hierarchical structure is equally important in the way that it helps to define what constitutes a possible segment by showing some of the constraints on possible feature combinations. A weakness of the older distinctive feature theories was the inadequacy of the formal devices for preventing impossible combinations of features. If there were even as few as 12 independent binary features, they could specify $2^{12} = 4096$ possible segments; given a larger number of features this soon becomes a ludicrously large number. Although, as we will see, much remains to be done, a feature hierarchy sets some limits on which feature values can co-occur.

The hierarchy in Figure 3.1 formalizes some very traditional phonetic concepts. The basic division into five hyper-features reflects the standard practice of articulatory phonetic description as seen in many textbooks. Abercrombie (1967), for example, notes that consonants can be described in terms of the place of articulation, the manner of articulation, the oro-nasal process, the state of the glottis, and the airstream mechanism. The same organization is apparent in Pike (1943), and has been taken over by Ladefoged (1971, 1982).

The first two hyper-features correspond to the concepts of place and manner, interpreted in a slightly different way, with the emphasis on the active articulator, and the allocation of places to sounds irrespective of whether they are consonants or vowels. The third hyper-feature, Nasality, is regarded by traditional IPA phoneticians as part of the manner of articulation (e.g. in the IPA consonant chart); but here, in accordance with the usual generative phonology tradition, Nasality is regarded as a separate process, so that, as with the other hyper-features, it is also applicable to consonants and vowels alike. When this is done it is possible to describe, for example, the nasalization of vowels before nasal consonants as involving the spreading of the feature value [+nasal]. The Laryngeal hyper-feature has long been recognized as a separate entity by all phoneticians, although the structures that it now dominates have more extensive scope. The final hyper-feature, Airstream Mechanism is often not considered as an independent node by generative phonologists, who usually regard the Laryngeal hyper-feature as specifying all the glottalic aspects of sounds, the Place and Manner (Stricture) hyper-features as specifying a velaric airstream when necessary, and the pulmonic airstream being sufficiently invariant to require no specification. As will become apparent later, there are advantages in setting up an additional hyper-feature in accordance with the IPA tradition.

3.2 Place of articulation

The Place hyper-feature has an additional level of structure before the terminal elements are specified. The IPA tradition, and that adopted in many textbooks, such as Ladefoged (1982), has been to regard place of articulation as a continuum from the glottis to the lips, with major points indicated, but with little or no structure defined among the points along the continuum. It now seems to me that this is wrong, and the arguments advanced by Halle (1983) and many others are correct in emphasizing the importance of the active articulator. In this way one can make more satisfactory statements concerning what double articulations are possible, and what articulators are free to make secondary articulations. Perhaps equally importantly, the notion of adjacent articulations can be given more meaning, if the description first indicates the active articulator. Thus a velar and a uvular stop are really adjacent articulations, differing only by a comparatively small movement of the body of the tongue; but a sublaminal retroflex and a palatal stop are far apart, one being a type of coronal articulation, and the other involving dorsal activity, making them very distinct sounds and not really 'adjacent' although they might be in adjacent columns on a consonant chart.

The major nodes of the hyper-feature place may not need to be further distinguished in many phonological statements. It is often sufficient to characterize a sound as being simply Labial, Coronal, Dorsal or Radical. In such usages these terms are acting as cover features, or unary features, a point to which we will return shortly. But although the phonologies of most languages do not need to make subtler distinctions, the phonological feature set has to provide finer distinctions, such as that between bilabial and labiodental sounds, because a few languages (e.g. Ewe and Venda) make use of them. Generally speaking we can use a set of default feature specification rules to fill in the lower level feature values, making it clear that, for example, Labial sounds are by default [bilabial] if they also have the value [stop], and by default [labiodental] if they also have the value [fricative]. A complete set of default feature specification rules will not be given here, as the hierarchy of features in Figure 3.1 has not yet been sufficiently tested. Interesting work on these lines for a more standard feature set has been done by Dogil (1988).

Labial articulations are of three different types. The first, which involves the vertical movement of of one lip towards the other, we will regard as the modal or neutral form of labial compression; this is the movement involved in a normal bilabial stop. The second is rounding, the drawing of the corners of the lips together, which is the movement involved in a labialvelar approximant [w] (This sound also involves a high back tongue gesture). A third possibility might be that the lips can be protruded without necessarily drawing the corners forward. But from a phonological point of view, rounding and protrusion do not seem to be distinct gestures; they may be combined in ways that are specific to individuals, or perhaps (as suggested by Linker 1982) in ways that are characteristic of languages; but we will not distinguish between them here. A truly different lip gesture is that involved in labiodental sounds, in which the lower lip is retracted so that it approaches the upper teeth. These three values correspond to what, in a different context, Keating (1984) calls major phonetic categories. They can be regarded as the three named values within the Labial continuum: [protruded], [compressed], and [retracted].

A constraint that should be made evident by the feature system is the fact that bilabials can be accompanied by lip protrusion, whereas labiodentals cannot. Labiodentals can be followed by a high back rounded (i.e. labial velar) approximant, and in this sense can be rounded; but while the lower lip is being retracted for a labiodental it cannot be simultaneously protruded. It is interesting in this respect to note the details of the lip movements that occur in some of the languages that contrast bilabial and labiodental fricatives. In Ewe (as shown by the photographs in Ladefoged, 1968), the bilabials [ɸ,β] are made with simple lip compression, whereas the labiodentals [f,v] are clearly distinguished by actively raising the upper lip, a gesture that makes them sound even further from any impression of lip rounding that might be associated with the bilabials. In Venda the reverse procedure occurs. The labiodentals [f,v] do not have a distinctive raising of the upper lip; but the bilabials [ɸ,β] do sometimes have a slight lip protrusion that emphasizes their non-retracted character, and makes them sound slightly rounded.

The notion of a continuum within the feature Labial is captured in Figure 3.1, by regarding Labial as a three valued feature, the three named possibilities being [protruded], [compressed], and [retracted]. These values are also sufficient to account for the differences in lip shape that distinguish the Swedish high front rounded vowels [y] and [ʉ]. Both these vowels have a value for Labial, [y] being [protruded], and [ʉ] being [compressed].

The phonological representation of labial gestures in terms of the three values [protruded], [compressed], and [retracted] involves a simplification of the phonetic facts. These three values are not really on a single dimension. The opposite of protruded lips is not a retracted lower lip; it is spread lips, in which *both* lips are retracted by drawing back the corners. Consequently it is a distortion to regard the formation of labiodentals as involving the opposite of protrusion; in forming labiodentals the lower lip is slightly retracted while the upper lip is not. Nonetheless the use of these three terms as oppositions within a single feature is completely appropriate. The question we must ask is whether it allows us to capture a phonological truth, or whether it simply obscures the phonetic facts. The answer is evident when we recall the purpose of these representations. We are trying to set up a system for representing just the phonological facts of languages. The simplification of the physiological phonetic facts allows us to see patterns of sounds that are otherwise not evident, and it is therefore fully justified. By regarding the spreading of the lips in a non-rounded vowel as equivalent to the retracting of the lower lip in the formation of the labiodental we can see a phonological universal: labiodentals cannot be contrastively protruded. If we had used two binary features, one to specify the difference between labials and labiodentals, and the other to specify that between sounds with lip protrusion and sounds with lip spreading, we would not have been able to make this point, except by an *ad hoc* statement. Putting this another way, we are making the claim that the differences between sounds made with spread lips and labiodentals are a matter of phonetic detail which does not play any role in phonological descriptions.

One final point of possible confusion concerning the specification of labial gestures needs to be cleared up. I have called the major place node Labial; and I have also called the feature it dominates Labial, the possible values of this feature being [protruded], [neutral], and [retracted]. As explained earlier, I capitalize the names of features, and use square brackets and lower case for the values that can occur. It might have been less confusing if I had called the major node Labiality, and reserved Labial for the name of the feature, but as contemporary usage seems to be to make no distinction between the major node and the feature in this case, I have retained the same names.

When we come to consider Coronal articulations, it becomes obvious that many points are still not clear, particularly with regard to the interactions between the tip, blade, and front of the tongue. As these terms themselves do not necessarily mean the same thing for all phoneticians, I had better begin by stating how I will use them. There is no way in which one can be precise about these matters both because the tongue is a continuous body with no relevant major anatomical landmark and because people vary considerably in the shapes of their mouths. I will use the terms as indicated in Figure 3.2.

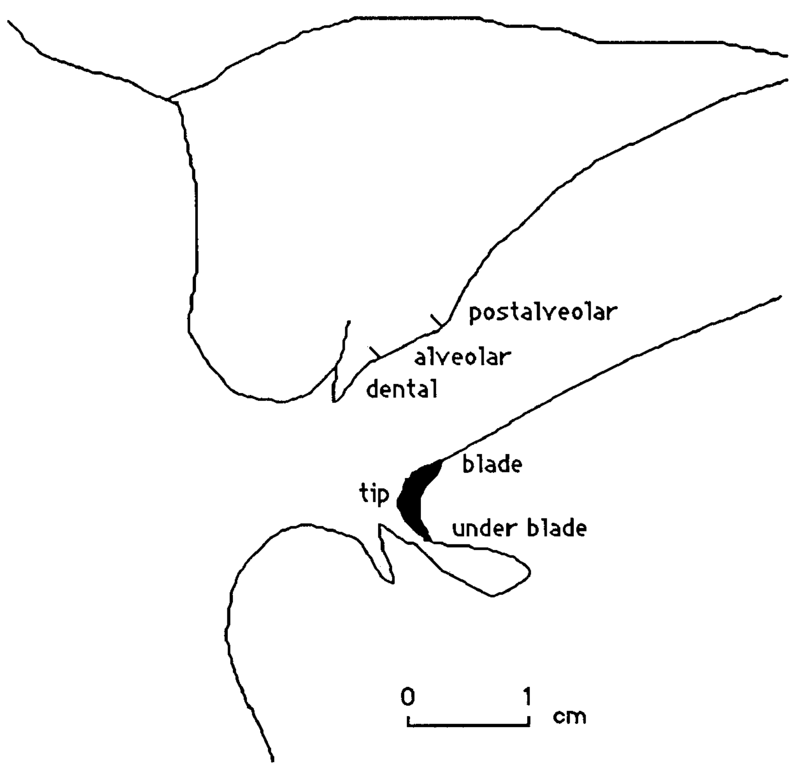

Figure 3.2. Definitions of terms used in defining the features dominated by the Coronal node.

The tip is considered to be just the part that has a primarily vertical aspect (i.e. largely parallel to the surfaces of the teeth) plus a small area about 2 mm wide on the upper surface. Sounds made with the tip of the tongue are said to be apical. The underside of the tip of the tongue is used in some articulations; sounds made in this way are said to be sublaminal. Behind the tip is the blade, which is the defining part of the tongue for sounds that are said to laminal. It is difficult to say how far back the blade extends before what is called the front of the tongue begins. One answer would be to say that it extends for 8-10 mm behind the tip, making the front begin 1 cm back from the most forward point, the very tip of the tongue. This would coincide in my case with the extent of the mobile part of the tongue not attached to the floor of the mouth (by the frenulum). But, because of people's different head sizes, and great differences in the extent of the frenulum, a more useful definition of the blade of the tongue from a linguistic phonetic point of view would be in terms of its relation to the roof of the mouth. The point defined above (about 1 cm from the most forward point, this being in my case, the limit of the freely moveable part) would also coincide for me with the point immediately below the center of the alveolar ridge when the tongue is at rest. This, of course, requires us to define the center of the alveolar ridge – an equally difficult task. I will take it to be the point of maximum slope behind the upper teeth in the curvature of the midline sagittal section of the vocal tract. In practice this is often difficult to determine but it is probably the most useful point that can be approximated in a wide selection of individuals. The best definition of laminal sounds seems to me that they are made with the part of the tongue which, when it is at rest, is in front of the center of the alveolar ridge.

The determination of the center of the alveolar ridge also enables us to define three parts of the upper surface of the vocal tract that are relevant to the discussion of Coronal articulations. Sounds made on the teeth are obviously dental; those that are made behind the teeth but in front of the center of the alveolar ridge are alveolar; and those that are made immediately behind the center of the alveolar ridge are postalveolar.

These distinctions allow us to define two terminal features dominated by the Coronal node: Apicality, which specifies the action of the lower, active, articulator; and Anteriority. which specifies the part of the upper, passive, articulator that is involved. Each of these features has three named values as shown in Table 3.1. This table also shows estimates of the likelihood of different combinations of the two features. It should be emphasized that these are simply guesses, based on my experience in investigating a large number of languages. As Maddieson (1984) notes, the published data on this matter are not very reliable. Many writers do not distinguish between by different languages; examples of contrasts in many of these languages are given in Ladefoged and Maddieson (1985). Of the nine possibilities shown, Toda is the only language that I know of that may contrast more than four; and I am not sure how accurately I have pushed the published description of Toda into the categories of my classificatory scheme. All the other languages listed in Table 3.2 reflect observations based on my own fieldwork, and I am more willing to stand by these descriptions. These data allow us to consider whether we really have evidence for saying that there are two three-valued features. It seems to me clear that more than two binary features are required. It would take hugely procrustean efforts to classify all the contrasts with only two binary features. If we are to retain the ability to compare the phonological description of one language with that of another, as suggested in Chapter 2, then I cannot think of any definitions that would enable us to account for these data in terms of two binary features, or even in terms of one binary feature and one three-valued feature.

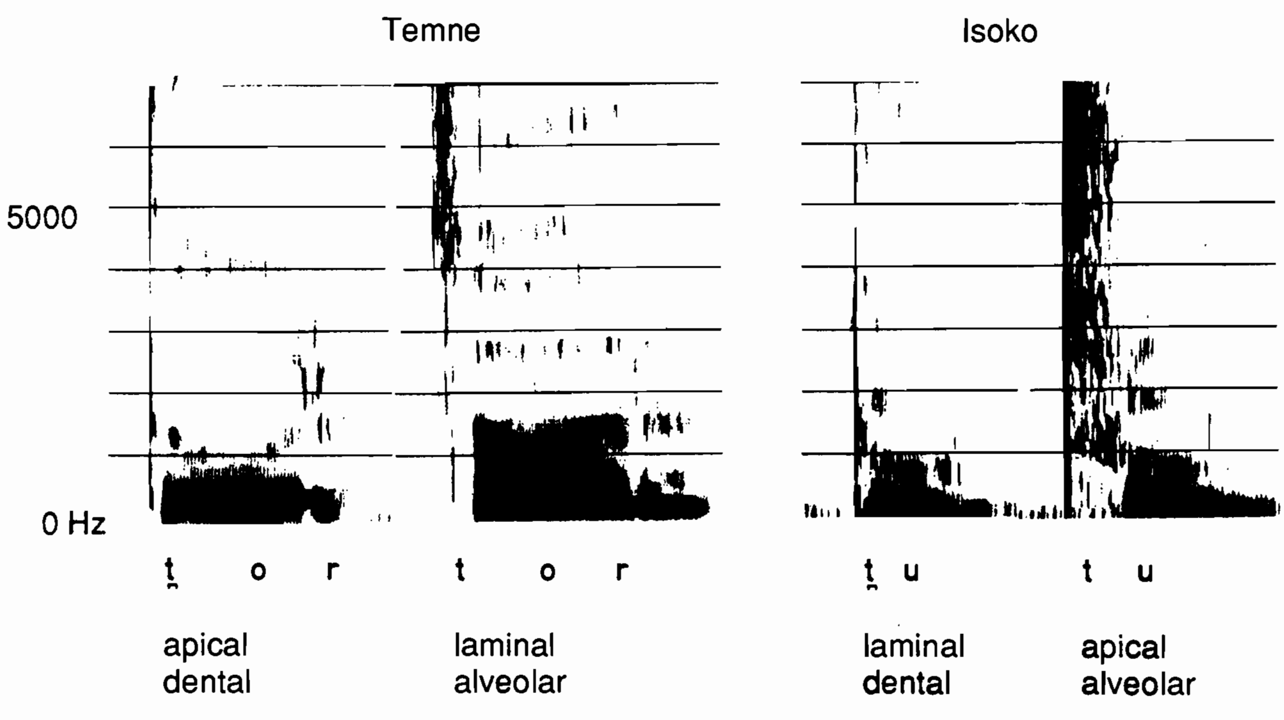

Figure 3.3 Spectrograms of laminal dental and apical alveolar voiceless stops in Isoko, and apical dental and laminal alveolar voiceless stops in Temne.

Any three-valued feature can, of course, always be replaced by two binary features. But, when this is done, it is no longer self evident that a segment cannot be dental, apical alveolar, sublaminal postalveolar and apical postalveolar stops. The first three of these combinations also occur contrastively in Malayalam. Consistent differences between the last two possibilities are known to occur between languages (Ladefoged and Bhaskararo 1983), and it is not at all improbable that these differences should be phonologically distinct within a single language such as Toda.

Apical postalveolar sounds are the most common forms of retroflex consonants, occurring not only in the Indo-Aryan languages of Northern India, but also in their in languages such as Ewe, in West Africa and many Australasian languages such as Nunggubuyu and Wangurri. Sublaminal postalveolar retroflex consonants seem to be mainly limited to Dravidian languages in the Southern part of India. Laminal postalveolar sounds (in IPA terms palatoalveolars) contrast with apical postalveolar sounds (IPA retroflex sounds) in Australasian languages and elsewhere.

Table 3.2 Languages illustrating the use made of different combinations of the possible values of the features Apicality and Anteriority.

		Anteriority	
	[dental]	[alveolar]	[postalveolar]
Apicality			
[laminal]	Malayalam Toda Nunggubuyu	Temne	Malayalam Toda Nunggubuyu
[apical]	Temne	Malayalam Toda Nunggubuyu	Hindi Toda Nunggubuyu
[sublaminal]			Malayalam Toda

What then can we say about the possible articulatory interactions betwen the tip, blade and front of the tongue as thus defined? Perhaps the best way of answering this question is by considering a number of articulatory contrasts that have been observed within languages. Table 3.2 shows the use made of some of the possibilities by different languages; examples of contrasts in many of these languages are given in Ladefoged and Maddieson (1985). Of the nine possibilities shown, Toda is the only language that I know of that may contrast more than four; and I am not sure how accurately I have pushed the published description of Toda into the categories of my classificatory scheme. All the other languages listed in Table 3.2 reflect observations based on my own fieldwork, and I am more willing to stand by these descriptions. These data allow us to consider whether we really have evidence for saying that there are two three-valued features. It seems to me clear that more than two binary features are required. It would take hugely procrustean efforts to classify all the contrasts with only two binary features. If we are to retain the ability to compare the phonological description of one language with that of another, as suggested in Chapter 2, then I cannot think of any definitions that would enable us to account for these data in terms of two binary features, or even in terms of one binary feature and one three-valued feature.

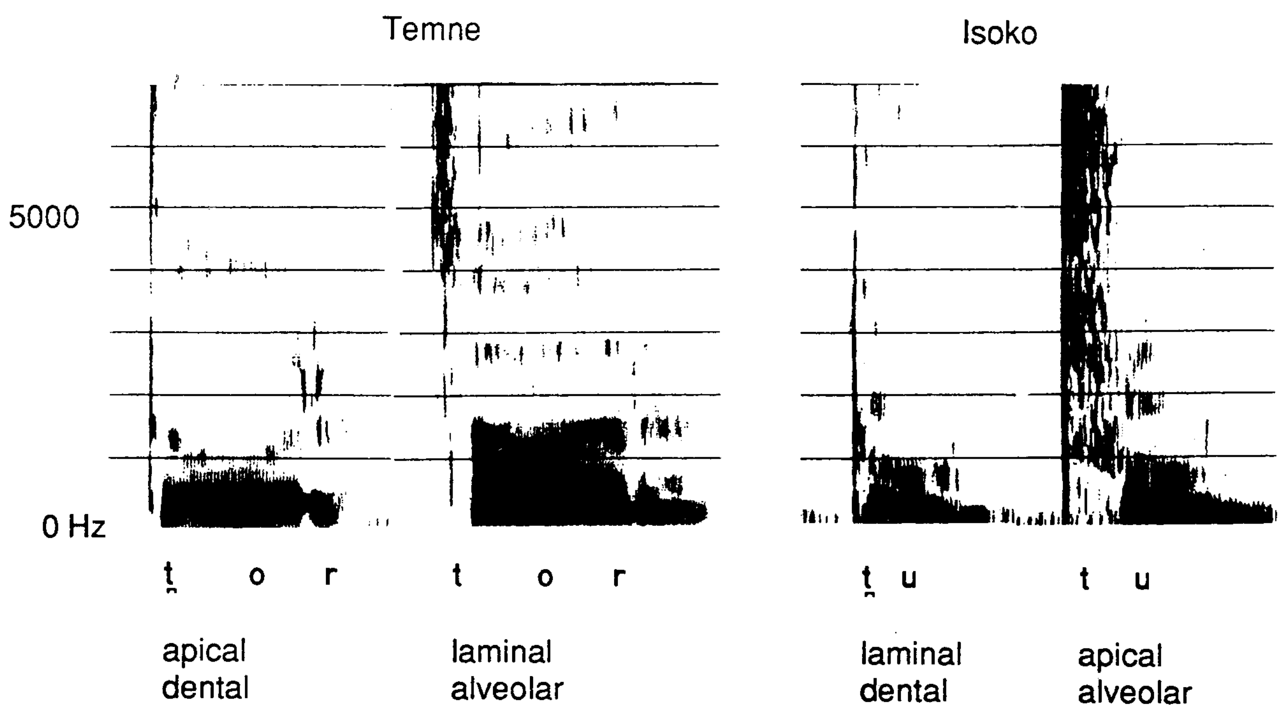

Figure 3.3 Spectrograms of laminal dental and apical alveolar voiceless stops in Isoko, and apical dental and laminal alveolar voiceless stops in Temne.

Any three-valued feature can, of course, always be replaced by two binary features. But, when this is done, it is no longer self evident that a segment cannot be be specified as being both [dental] and [postalveolar]. Furthermore, as was also discussed in the preceding chapter, this will remove the possibility of correctly stating some phonological observations. For example, the use of three-valued features in Table 3.2. shows how Malayalam uses the three way possibilities along the diagonal between Apicality and Anteriority to maximize contrasts among similar sounds, a fact that would not be evident if four binary features were used to classify these sounds, as suggested by Mohannan and Mohannan (1984). Furthermore the three way division of the feature Anteriority offers an appropriate way of showing the low level allophonic variations that occur in English words such as 'eighth, eight, tray' which in many pronunciations have dental, alveolar, and post alveolar allophones of /t/.

All this having been said, I will freely admit that neither Anteriority nor Apicality presents a strong basis for arguing the case for multivalued features. Generally speaking, phonological processes involving these two features can be adequately expressed using binary values. I will retain the three values shown in Figure 3.1 and Tables 3.1 and 3.2, because, once we have allowed the possibility of multivalued features (as we surely must for features such as Height), there are the slight advantages mentioned above in allowing the three possibilities. In addition, the three by three matrix used in Tables 3.1 and 3.2 very properly forces us whether there could be languages that use sublaminal dental or sublaminal alveolar sounds. The theory proposed here suggests that there could be such languages, and the fact that they have not been attested is the result of an accidental gap in the inventory of known phonological structures. In this way we can retain the notion that one of the functions of a feature theory is to provide an account of human linguistic phonetic capabilities.

We noted earlier that the feature Apicality refers to properties of the active articulator, the tip or blade of the tongue; whereas Anteriority refers to the passive articulator, the site on the roof of the mouth of the articulation. As both features are dominated by the major node Coronal, it is a constraint of the theory that one of the three sites of the passive articulator can be specified only when one of the three possibilities for the active articulator is also specifiable. Gorecka (personal communication; forthcoming MIT PhD thesis) has a more far reaching proposal, which is tantamount to reviving in a fuller form the traditional notions of active and passive articulators. She proposes a feature hierarchy in which there are two sister nodes, Site (= passive articulator) and Articulator (= active articulator), instead of the hyper feature Place. This seems to me to be a retrograde step, in that the theory would then permit a great number of combinations of Site and Articulator that never occur, making it possible, for example, to specify the equivalent of apical pharyngeal sounds.

It is, of course, true that our linguistic theories exist alongside other theories, and we could simply say that impossible articulations are ruled out by physiological constraints that are not part of a linguistic theory. But it is not clear why these particular physiological constraints should be permitted to be outside linguistics. We have already relied on a physiological basis for other parts of the theory. When we say that there are four and only four types of articulatory gestures (labial, coronal, dorsal and radical) we are encapsulating within our linguistic theory some of the physiological constraints on possible articulations. A feature theory should, in itself, provide an account of what is, and what is not, linguistically possible. It is part of linguistics, not physiology, that no language has sublaminal uvular stops. Accordingly we should have a feature theory with a constrained notion of possible types of Coronal and other articulations.

This is a good moment to make clear another matter concerning the kind of feature specification proposed here. If a sound is specified as being Coronal it must have potential values for both the feature Apicality and the feature Anteriority. They may be plus or minus if binary oppositions are being maintained, or [laminal], or [apical] or [sublaminal] and [dental], or [alveolar] or [postalveolar] in a three valued system. As we noted earlier, the values may be specified by default whenever Coronal is specified. But it is not possible to have a sound that is Coronal and does not have potential values for both Apicality and Anteriority. It is only when a sound is, say, Labial and not Coronal, that neither Apicality nor Anteriority is specifiable. Thus at an abstract level in the phonology, the major place nodes may act as cover features, designating sets of values of the terminal features. But the major nodes are not themselves features. If Coronal were to be a feature then it would have to be defined in terms of some properties; and I cannot see how this is to be done independently of the characteristics of Anteriority and Apicality or (Distributed In the SPE system).

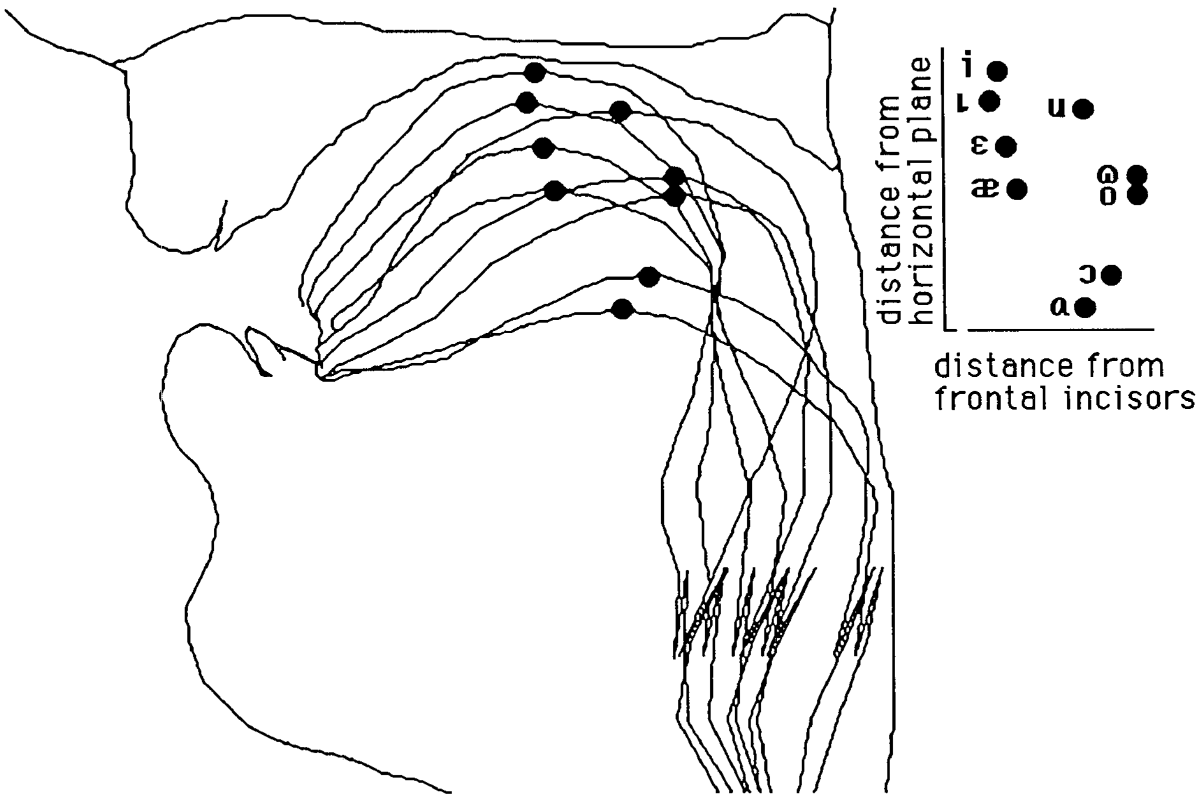

Figure 3.4 The mean tongue positions in American English vowels. The points show the locations of the highest point of the tongue with reference to the tips of the frontal incisors and a horizontal plane parallel to the upper molars.

The next major node, Dorsal, dominates the features necessary for specifying sounds involving the body of the tongue. This node, differs from the Labial and Coronal nodes we have been discussing by virtue of its major role in classifying both consonants and vowels. For some vowels , one aspect, the degree of lip rounding, is specified by a feature dominated by the Labial node; and for a few rare vowels (e.g. American English and Chinese r-colored vowels) the role of the tip or blade may require specification of one or more Coronal features. But nearly all vowels can be described in terms of one or more dorsal features (in addition to the auditory features Height and Brightness discussed in the previous chapter). I have retained the terms High (Low) and Front (Back) for the articulatory features as shown in Figure 3.1, although it is not at all clear that the classes of vowels defined by tongue body positions are the same as those defined by the traditional use of these terms. The mean position of the tongue body in American English vowels as produced by five speakers is as shown in Figure 3.4 (based on data and calculation of means by factor analysis as described in Harshman, Ladefoged and Goldstein 1977). As can be seen, the tongue height of [u] is below that of [ɩ], and the tongue height of [o] is below that of [æ]. The traditional front-back descriptions are also not in accord with the actual articulatory data in Figure 3.4. Diagrams showing the height of the tongue body (as advocated by Fischer-Jørgensen, 1985) would reveal much the same relationships.

There are a number of alternative possible ways of quantifying the position of the tongue in vowels. The most well known is that used by Stevens and House (1955) and Fant (1962), who point out that, from an acoustic phonetic point of view, the most important articulatory characteristics of vowels are the position of the point of maximum constriction of the vocal tract, and the cross sectional area of the vocal tract at that point. Figure 3.5 shows the vowels in Figure 3.4 arranged in this way. The groupings in this figure do not form any obvious natural classes from a linguistic point of view.

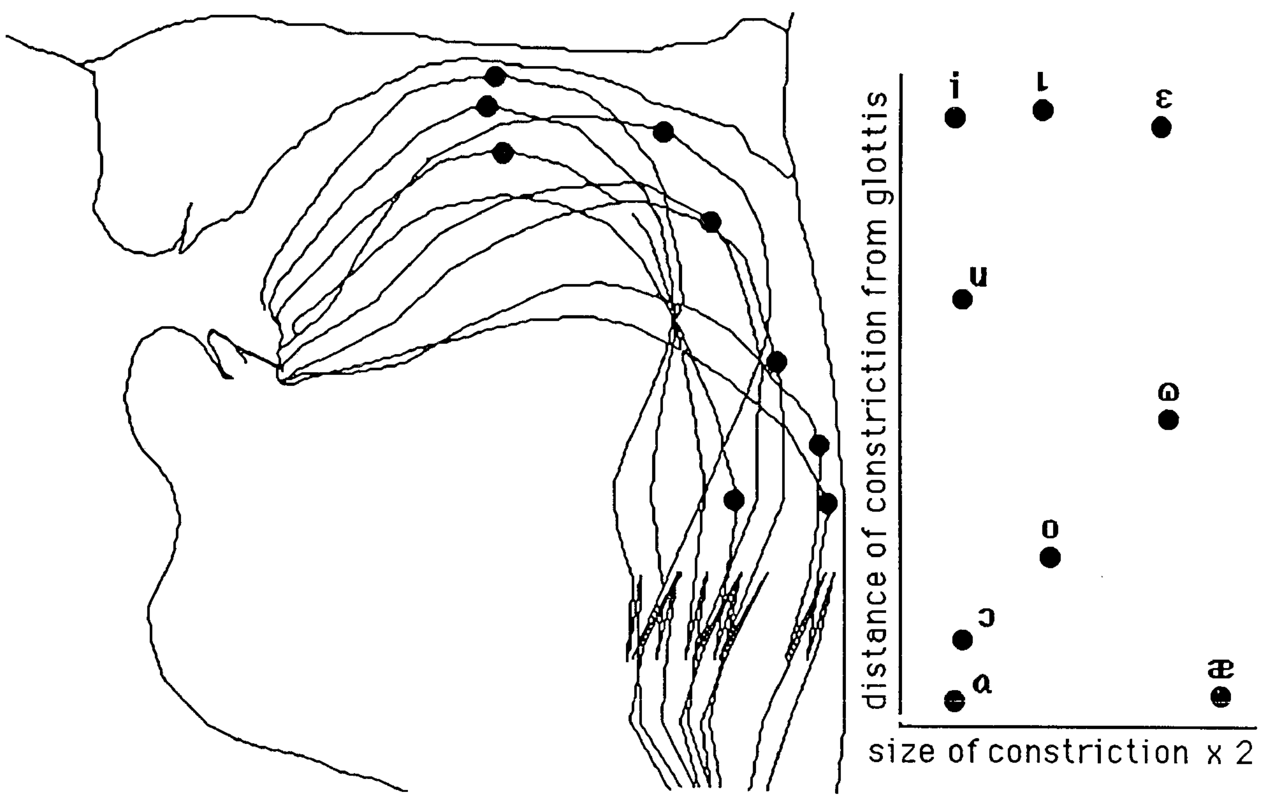

Figure 3.5 The mean tongue positions in American English vowels measured in terms of the size of the constriction, and the distance of the constriction from the glottis, as suggested by Stevens and House (1955) and Fant (1962).

A compromise between the two characterizations we have considered so far has been suggested to me by my colleague Ian Maddieson. In this view each vowel is characterized in terms of the distance of the highest point of the tongue from the roof of the mouth, as shown in Figure 3.6. For the front vowels below the hard palate this is effectively the same as the position of the highest point of the tongue as in Figure 3.4. For the back vowels it is somewhat different, especially as it must be remembered that the height of the soft palate is directly correlated with the height of the vowel, so that the low vowels are in fact closer to the roof of the mouth than appears in the figure, which shows only one position of the soft palate.

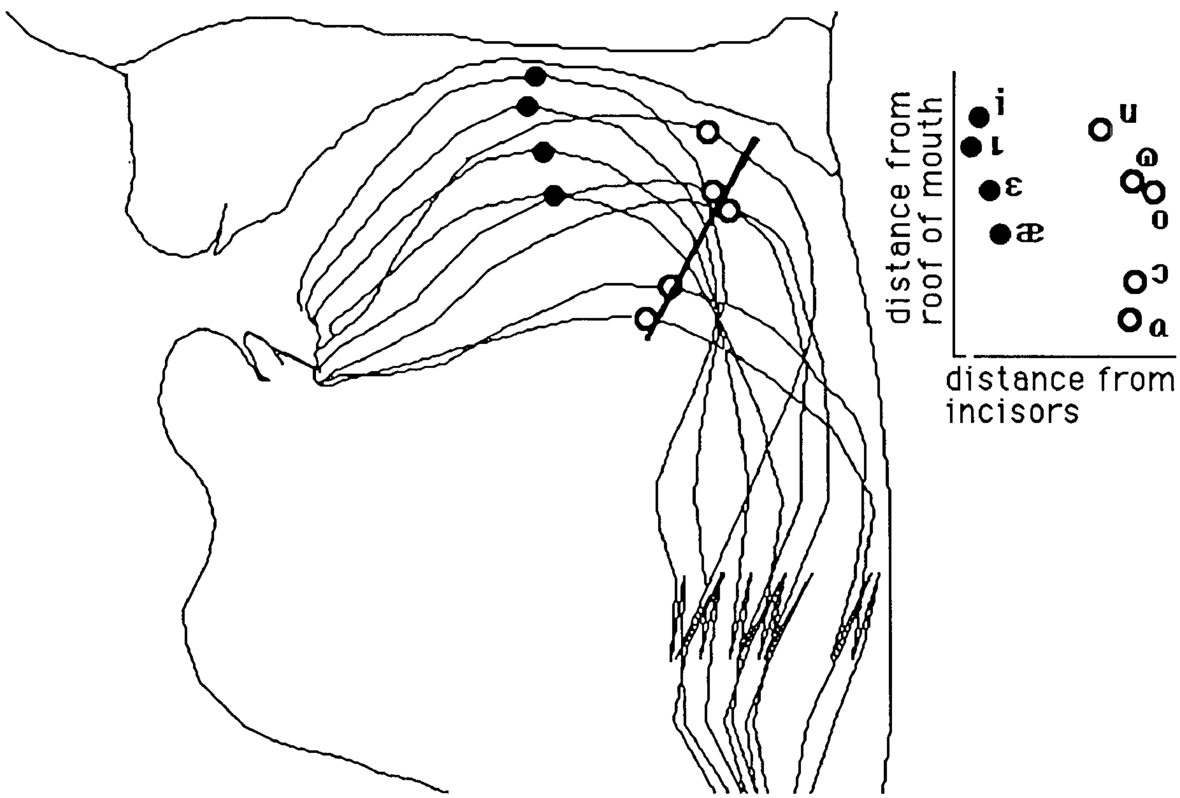

Figure 3.6 The mean tongue positions in American English vowels:measured in terms of the distance from the roof of the mouth. The lowering of the soft palate that takes place in low vowels has been taken into account in the placement of the points on the right of the figure.

However, even this representation of the articulation of vowels is not as close to the usual linguistic representation of these vowels as is that provided by the acoustic data. Figure 3.7 compares the mean formant frequencies of these same vowels with the articulatory representation of figure 3.6. The acoustic representation corresponds more closely to the usual linguistic description. Note, for example, the way in which the low vowels have approximately the same height, and the vowels [u] and [ʊ] are slightly forward as they are in American English. Either the vowel [ɪ] or the vowel [ɛ] seems not quite in the place in which a phonetician might have plotted it, but even these discrepancies are minor.

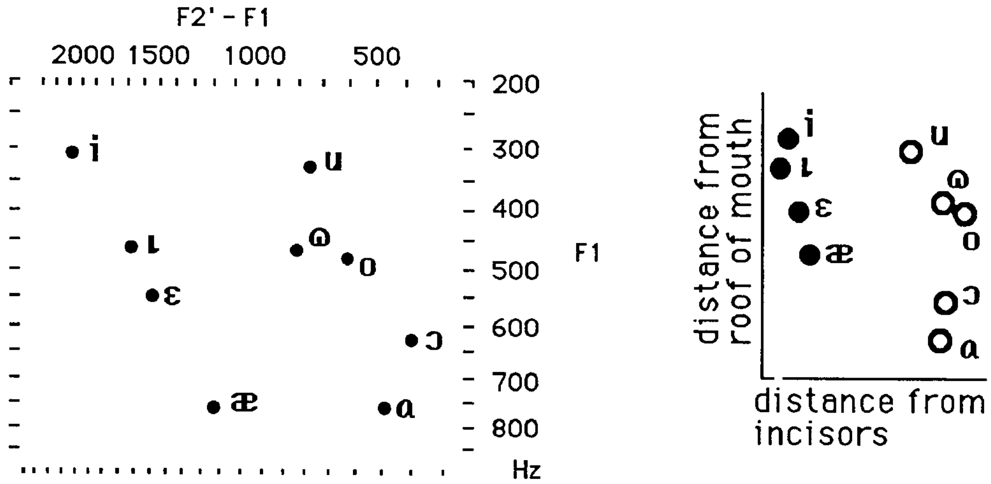

Figure 3.7 A comparison of an acoustic representation and an articulatory representation of a mean set of American English vowels.

The interactions between consonants and vowels often involve features dominated by the Dorsal mode and we need feature specifications that link vowels and consonants appropriately. By using the same set of articulatory features for both vowels and consonants we can make it clear that both these classes of sounds have a great effect on one another precisely because they are produced within the same mouth. We need a phonetic theory that allows us to explain why velar consonants are more likely to have an advanced articulation before high front vowels than before back vowels, as they might well if there were no connection between the articulation of vowels and that of consonants. The same theory should make it evident that high vowels are more likely to be lowered by uvulars (as occurs in Serer, French, and many other languages) than for the reverse to occur so that mid vowels are raised by uvulars (which, to the best of my knowledge, never happens). The feature set proposed here allows explanatory statements of this kind. The features High and Front are multivalued features , each describing an ordered set of possibilities. In this way high back vowels may be associated with velar consonants, mid back vowels with uvular consonants and low back vowels with pharnygeal consonants. For front and central vpwels the associations with consonants are less clear. High front vowels are akin to palatals. But mid front and central vowels have less consonant-like properties, and low front vowels cannot be said to be associated with any particular types of consonants. From a phonological point of view, they may be left unspecified with respect to these features, using the Height and Backness features discussed in the previous chapter to account for interactions among vowels.

True palatal consonants requiring the feature specification [front] are rare. Many of the sounds that are called palatals are more properly categorized as laminal postalveolars (for example, French [ɲ] as in 'agneau' and Italian [ʎ] as in 'figlio'). But I have heard true palatals, that is, sounds in which the body of the tongue is raised towards the hard palate well behind the postalveolar region, in, for example, Ngwo (see palatographic data in Ladefoged 1968, reprinted in Ladefoged and Maddieson 1986). Mohannan and Mohannan (1984) report that in one dialect of Malayalam there is a contrast between palatoalveolar (i.e. laminal postalveolar) nasals and true palatal nasals. Ladefoged and Maddieson (1986) cite Jaqaru (Hardman 1966) as giving clear evidence of contrasts between true palatal, velar, and uvular stops.

If we take it as axiomatic that more complex segments are less likely to occur than simple segments, then we can account for the comparative rarity of palatals by following a suggestion made by Keating (personal communication), and considering them to be complex segments. They are complex in that they have both Coronal and Dorsal attributes, having both the blade and the front of the tongue raised. This solution is in accord with palatographic data such as that of Doke (1931) for Shona [ɲ]. But it does not agree with the palatographic data for Ngwo, nor with the x-ray data for Hungarian palatals in Bolla (1980), both these sources show no Coronal articulation (the tongue blade is not raised) for the palatal stops. The feature system in Figure 3.1 allows palatals to be classified as either [high, front] Dorsal sounds or as complex segments that are both Coronal and Dorsal.

The final major place node Radical, classifies articulatory gestures made with the root of the tongue. Unlike the other three place nodes Radical dominates only one feature, here called in accordance with recent tradition, Advanced Tongue Root (ATR). This feature is relevant in the descriptions of comparatively few languages. Furthermore, it seems that phonologies of even these languages require only two values, which may be referred to as [+ atr] as opposed to [- atr], or [advanced tongue root] as opposed to [epiglottal]. As was first shown by Ladefoged (1968), the vowel harmony sets in languages such as Igbo can be described in terms of the advancement of the tongue root; in these languages it is more usual to use the specifications [+ atr] and [- atr]. This feature is also relevant in the specification of so-called pharyngeal consonants in Arabic and other Semitic languages, although in these cases it is the value [epiglottal] that is more appropriately specified. The strident vowels of some Khoisan languages (Snyman 1975, Traill 1985, Ladefoged and Traill 1980) can also be classified as [epiglottal].

There are a number of sounds that have two active articulators, so that they form complex segments characterized by segments dominated by two major nodes. One of the justifications for the division of the traditional places of articulation into four groups, each dominated by a single major node, is that this framework allows us to specify all and only the combinations that can occur. As has been noted by Sagey (1986), complex segments formed by combinations of the major place nodes within a single segment are not uncommon. Labial plus Dorsal articulations as in [kp, gb] are the best known; Coronal plus Dorsal articulations occur in clicks; and Radical plus Dorsal articulations occur in some Caucasian fricatives (Catford, 1977b). We have also noted that palatal sounds can be considered as complex segments with Coronal and Dorsal attributes. Furthermore, what are traditionally known as secondary articulations (labialization, palatalization, velarization, pharyngealization) can be regarded as combinations of two different places involving different strictures.

3.3 Manners of Articulation

More work is needed in the characterization of the hyper-feature Stricture. The hierarchical structure of the features dominated by this node is extremely hard to formalize, but it seems clear that the foremost manner of articulation is that determining the degree of Aperture. Following the SPE usage, many contemporary phonologists maintain that there are two important articulatory manner features, one specifying the opposition Sonorant - Obstruent, and the other the opposition Continuant - Noncontinuant (stop). Neither of these pairs of oppositions is defined in the traditional way by Chomsky and Halle (1968). The term Sonorant originally meant just voiced non-obstruent sounds, a meaning which is retained by the definition of the auditory feature given in the previous chapter. In SPE usage sonorants also include [h,ʔ] and voiceless nasals and laterals, providing a natural class that I do not find as useful as the classical grouping. The other differences from classical usage are in the term stop, which used to mean a complete stoppage of the airstream, and therefore excluded nasals, and the term continuant, which used to mean a sound that could be continued, and therefore included nasals. In contemporary usage nasals are regarded as stops. This is a comparatively minor terminological change, which I will follow, as it leads to useful phonological groupings.

A more real problem with contemporary feature theories is that there is no simple way of defining the natural class of fricatives such as [f,θ,s,ʃ]. It is, of course, not sufficient to include a feature Sibilant (Strident) as suggested in Chapter 2, because not all fricatives are sibilants. It is more appropriate to set up an articulatory feature, Aperture, with [fricative] as one of its named values as shown in Table 3.1. There are many advantages to having a feature Aperture with the three possibilities [stop, fricative, approximant]. These items form a set of mutually exclusive possibilities; and, as we have been noting, it is an important aspect of any phonology that it should be able to make clear what can and what cannot co-occur. Furthermore, as I noted earlier (Ladefoged 1971:55): "These values form a linearly ordered set, by means of which we [can] give an explanatory account of lenition phenomena, in which stops weaken to fricatives, and a further weakening gives rise to approximants." This phenomenom is exemplified in the pronunciation of the Danish word "Ladefoged" (barn steward). The original voiced stops in this word have weakened, first to fricatives [ð] and [ɣ], and then to approximants (frictionless continuants) for which there are no distinct IPA symbols, eventually becoming (at least in the case of the velar consonant) omitted altogether in modern Danish. Progressive changes from stop through fricative to approximant are not easy to explain in terms of binary features. In the system being proposed here the Danish changes (and similar variations in other languages such as Spanish) are simply assimilations in which the intervocalic stops become more and more like the surrounding approximants.

Previously (Ladefoged 1971:55) I rejected the notion of a single feature with the values [stop, fricative, approximant] because it did not permit fricative to be regarded as a separate feature that could be added to stops for the characterization of affricates. Now, however, formalisms have been developed (Goldsmith 1979, Halle and Vergnaud 1980) that allow us to note sequential attributes of a segment within a single slot in the CV structure. In this way affricates can be assigned the correct feature values in the correct order.

It is at this point that we should note a manner of articulation that has been deliberately left out of Figure 3.1: there is no mention there of taps or flaps. Following Inouye (1988), I will regard them as articulations that occupy a single slot in the phonological structure, but which have a sequence of feature specifications. Inouye points out that the essence of a tap is that it starts and finishes in an approximant position, and includes a brief stop in the middle. It is impossible to make a tap or flap starting from a [fricative] or [stop] value of the Aperture feature without inserting an epenthetic segment with an [approximant] value of this feature. As the necessary values of the Aperture feature are already available, there is no need to set up a special feature to take care of tap and flap articulations. They are, like affricates, dynamic segments in which different moments in the articulation have have different feature specifications. Inouye also points out a number of interesting phonological implications of this analysis, such as the simplifications that result in the formulation of the rules for the alternations in American English such as "fat, fatty".

We should no consider whether the distinction between a tap and a flap is worth pursuing. I noted earlier (Ladefoged 1971) that "A flap is ... distinguished from a tap by having one articulator strike another in passing while on its way back to its rest position, as opposed to striking immediately after leaving its rest position [in a tap]." It now seems to me that this may be only an incidental difference between taps and flaps, because flaps (if defined as in the quoted sentence) always have a more retracted articulation than taps. It may therefore be appropriate to consider a flap as a tap with a different place specification. I can see no phonological reason for distinguishing between the two. However, I am still a little uncomfortable with this, as the dynamics of the two gestures are so very different. A tap involves a small movement of the tip of the tongue, whereas a flap necessitates a backward movement of the body of the tongue, as well as a large upward and backward movement of the tip. A tap has the same transition into and out of it; a flap does not.

The next feature listed in Figure 3.1 is Trill. This is another feature that needs more investigation. It may be proper to consider trills as dynamic gestures much as suggested by Inouye (1988) for taps and flaps. But for the moment I will simply note the necessity for a feature of the kind that can specify Labial, Coronal and Dorsal trills in segments such as those described by Ladefoged, Cochrane and Disner (1977).

The next division among manner features provides the distinction between central and lateral sounds. Different values of the feature Lateral can occur with each of the values of Aperture. Distinctions between central and laterally released stops are common (e.g. in Mayan languages). Such differences could be regarded as sequential attributes of a single segment, as is appropriate for affricates. But it is less clear that this solution can be adopted in the case of clicks, which are also a form of stops that utilize the central-lateral opposition. Zhu|oãsi, for example has [k!, k!h, k||, k||h] , thus contrasting a central alveolar click accompanied by a voiceless unaspirated velar stop, a central alveolar click accompanied by an aspirated velar stop, and a pair of lateral alveolar clicks similarly accompanied. Simultaneous [+lateral] and [fricative] specifications are definitely necessary in the case of [s] and [ɬ] which contrast in Zulu and Welsh. Central and lateral approximants such as [ɹ] and [l] contrast in many languages, including most forms of English.

3.4 Nasality

The hyper-feature Nasality is fairly straightforward. It dominates a single terminal feature, Nasal, which applies to both consonants and vowels. From a phonological point of view, this is clearly a binary feature, in that sounds are either [+ nasal] or [- nasal]. I have heard only one language, Chinantec (Merrifield 1963), in which there are surface contrasts between oral, lightly nasalized and heavily nasalized sounds (Ladefoged 1971). These Chinantec sounds can be analyzed as having different numbers of segments in the underlying forms, the contrasts being of the form /a - ã - an/, the final consonant not appearing in the phonetic output. We must also note that there are severe co-occurrence constraints on this feature; among consonants only stops can be contrastively [+nasal] or [-nasal]. As will be suggested later, these kind of constraints will have to be handled with some notions of co-indexing that have not yet been worked out.

3.5 Laryngeal features

Before we look at the terms dominated by the Laryngeal node, we must consider the phonological relations that we want these features to capture. Virtually all phonologists regard the opposition voiced-voiceless as an important aspect of natural classes of sounds. A notable exception is Halle, who for many years (Halle and Stevens, 1971, Halle 1988) has advocated feature systems that replace the functions of the feature Voiced with various combinations of the four features Stiff, Slack, Spread and Constricted. Another widely recognized set of phonological possibilities is that for variations in Voice Onset Time (VOT). Although there may be low level phonetic differences, from a phonological point of view there are only three possibilities: aspirated, voiceless unaspirated, and voiced (Keating 1984).

We also need to be able to categorize differences in phonation types. Again, from a phonological point of view we need recognize only three types of voicing, breathy voice, modal voice and creaky voice (Ladefoged, Maddieson and Jackson 1988). Furthermore, although phoneticians (Laver 1980, Catford 1977a) have shown how three terms similar to these may be combinable from the phonetic point of view, the three named terms form a set of phonological mutually exclusive possibilities. There is never any need to classify a segment as being phonologically both [breathy] and [creaky]. Finally we must not overlook the obvious fact that specifications of tone and intonation can be regarded as part of the hyper-feature Laryngeal although I will not attempt to give here a system for the representation of these aspects of language.

In setting up a feature hierachy we must recognize that breathy voiced sounds are in some sense between voiceless and voiced sounds. In this way we can regard such things as the breathy voiced quality of the /h/ in English 'ahead' as being due to an assimilatory process. There is also the status of glottal stops to consider. They are, of course, voiceless sounds; and they are more likely to alternate with creaky voiced sounds than with voiced or breathy voiced sounds. Allophonic variations between glottal stops and creaky voice occur in Mazatec (Kirk, Ladefoged and Ladefoged 1984), Hausa (Lindau 1984) and many other languages. If creaky voice is in some sense between voiced sounds and glottal stops, this use of creaky voice can also be regarded as due to an assimilatory process. These factors point to there being an ordered set of five mutually exclusive possibilities; spread, breathy, modal, creaky and closed vocal cords.

It would be appropriate also to consider other interactions among laryngeal gestures, such as the alternations between voiced plosives and implosives that occur in widely different languages such as Thai and Zulu. We also need to note the well known connection between voiceless sounds and high pitch. But there are also converse facts to be acknowledged. Fully voiced stops (as well as voiceless unaspirated and aspirated stops) contrast with both voiced and voiceless implosives in Owerri Igbo (Ladefoged, Williamson, Elugbe and Uwulaka, 1976), making it impossible to equate full voicing with implosion. In most tone languages both high and low tones occur in syllables with both voiced and voiceless consonants, so we do not want a complete identification of high tone with voicelessness.

Bearing in mind that we want a feature theory that illuminates phonological universals, it seems appropriate to set up a somewhat redundant feature set, as shown in Figure 3.1. In this set there are four terminal features dominated by the Laryngeal node: Voice, specifying whether the vocal cords are vibrating or not; Glottal Stricture, indicating the degree of opening of the glottis; Aspiration, controlling VOT; and Pitch, which will not be considered further here. There are five named values for the different phonation types. These are all clearly mutually exclusive possibilities; if a sound has breathy voice, as in murmured sounds, it cannot be phonologically creaky, and so on. It may be that the five named types of glottal stricture can be regarded as the product of two ternary oppositions, as suggested by Stevens (1988). It is certainly appropriate to consider these glottal states as resulting from two physiological attributes of the vocal cords, their stiffness and their aperture. However there seems to be no phonological gain in doing this.

Note that the the feature Glottal Stricture does not include the terms [voiced] and [voiceless] as named values. Instead there is a separate feature, Voice, which specifies whether the vocal cords are vibrating, as in breathy, modal, and creaky sounds, or voiceless, as in both spread and closed types of glottal stricture. Just as the articulatory features High (Low), Back and Round do not of themselves explain why vowel systems are as they are, so too the feature Glottal Stricture does not provide a direct way of explaining why most sounds are either voiced or voiceless. There has to be a separate feature accounting for these two very natural classes of sounds. As we noted in the previous chapter, this feature could be given either an auditory or a physiological definition. At the moment it seems that both sets of properties distinguish the same classes of sounds, although further phonological evidence may later be forthcoming to show that one or other of these definitions provides slightly better groupings. Irrespective of whether it is considered to be an auditory or a physiological feature, there is no doubt that the feature Voice is a very necessary determiner of phonological classes.

Recent data (Dixit 1987a,b) indicates that the degree of aspiration is not simply related to the degree of glottal stricutre, as was once thought (Kim 1970, Ladefoged 1971). Furthermore, despite the fact that many languages have sets of stops with the three possibilities, voiced, voiceless unaspirated, aspirated (as in, e.g. Thai and Eastern Armenian), Aspiration is best regarded as a binary feature, with only the possible values [+ aspiration] and [- aspiration]. I cannot find convincing examples that show that a three term series is necessary for phonological reasons. There would need to be a rule whereby, for example, voiced stops become voiceless unaspirated stops in the same circumstances as voiceless unaspirated stops become aspirated (i.e. [b -> p] and [p -> p^h] in the same circumstances), Failing the occurrence of such a rule, and faced with other phonological facts in which aspirated forms of both voiced and voiceless stops (i.e. [b^h, p^h] etc) act together to form a natural class, it is better to consider Aspiration to be a binary feature. This allows us to formulate a simple explanatory statement of, for example, Grassman's law whereby aspirated stops of either kind become unaspirated if there is an aspirated stop of either kind in the word.

3.6 Airstream mechanisms

The final hyper-feature shown in Figure 3.1 is Airstream. Most contemporary feature sets do not provide an explicit way of characterizing just the sounds made with one of the three possible airstream mechanisms. They regard clicks, along with other velar sounds, as being specified by Dorsal attributes, and/or, along with implosives, by an ill-defined Suction feature; and they regard sounds made with a glottalic airstream mechanism as sufficiently specified by Laryngeal features. In the feature set proposed here, there is more redundancy. The Dorsal specification of clicks and the Laryngeal specification of ejectives and implosives is permitted; but there is also provision for each of these groups of sounds to be specified as a natural class of its own.

The hyper-feature Airstream dominates three features: Pulmonic, Glottalic, Velaric. It might seem as if there were no need to specify the presence of the pulmonic airstream mechanism, as it is present in all sounds; even clicks and ejectives still have a positive subglottal pressure. It is, however, necessary to note that some sounds have an increase in lung power associated with them. For example, Dart (1987) has shown quite conclusively that Korean so-called fortis stops have a significant increase in pulmonic pressure, as well as differences in phonation type.

I have called the possible classificatory values for the Pulmonic feature [+ fortis] and [- fortis]. More apt descriptive names might have been [+ heightened subglottal pressure] and [- heightened subglottal pressure], a pair of terms suggested by Chomsky and Halle (1968). I have avoided using these terms here, because they have been used to describe other phenomena, such as the opposition between aspirated and unaspirated sounds, a usage for which there is no physiological foundation (see Ladefoged 1971: 96). I have also rejected the possibility of using the terms [+ stress] and [- stress], although the most common manifestation of stress is an increase in subglottal pressure much of the kind used in Korean fortis stops. From a phonological point of view stress is usually a property of a larger unit than a segment, and therefore not like the features we have been considering. There are very few occasions on which it is appropriate to describe a consonantal element of the CV structure as having a value of the feature Pulmonic; segments such as the Korean fortis stops seem to be very rare.

Both the non-pulmonic airstream mechanisms occur in conjunction with the pulmonic mechanism (and sometimes, as in !Xóõ, in conjunction with each other as well). It is for this reason that each of the Airstream features has to be regarded as a major node, as well as a terminal feature. The glottalic airsteam mechanism has two mutually exclusive possibilities: ejective and implosive. Voiceless ejectives occur in many languages, for example Amharic, Navaho and Xhosa; voiced ejectives are unknown. Voiced implosives are many times more common than voiceless implosives, occuring in, for example, Sindhi and Kalabari. Both voiced and voiceless implosives occur at the same place of articulation in Owerri Igbo (Ladefoged, Williamson, Elugbe and Uwalaka, 1976). The likelihood of the different cominbatory possibilities is not adequately accounted for in the feature hierarchy in Figure 3.1.

The feature Velaric, with its two possible values [+ click] and [- click] provides the natural classes required in the description of Khoisan and Nguni languages. It would be possible to classify these sounds simply as examples of multiple articulations, as suggested by Sagey (1986); but there seems to be no phonological advantage in doing this, other than an apparent economy of description.

3.7 Hierarchical constraints

At the beginning of this chapter we noted that one of the uses of a feature hierarchy is that it constrains the theory by showing what can and what cannot occur. Most phonological descriptions do not require the specification of all possible features. But if our phonological theory is to be able to determine human phonetic capabilities, it must be able to show what kinds of full phonological specifications are possible, and what are not. In general, the full specification of every sound will involve passing through every hyper-feature node. Every sound must have the potentiality for having some Place, Stricture, Nasality, Laryngeal and Airstream specification. Possible exceptions are segments such as [h] and [ʔ], which may not require supralaryngeal specification. If this were so, it might be better to increase the depth of the hierarchy, as suggested in the original work on this topic by Clements (1985), who argued for a Supralaryngeal node that dominated all the hyper-features other than Laryngeal. I have resisted doing this because it is quite clear that it must be possible to specify Laryngeal sounds for Aperture. We need to be able to specify the degree of Aperture in glottal stops, for example, so that we can show that they have something in common with other stops in languages such as Hawaiian, where they belong in the same natural class as bilabial and velar stops.

At the next level of specification there are two possibilities: the hyper-features in Figure 3.1 dominate either major nodes or terminal features. There are two sets of major nodes, one dominated by the Place hyper-feature, and the other by the Airstream hyper-feature. Both sets of major nodes differ from hyper-features in that, instead of each major node having to be taken into account for the full phonetic specification of a sound, only one item is required from each of the two sets of major nodes as a minimum, although more than one may be specified. Features differ from either of the other two categories (hyper-features and major nodes) in that each feature that is dominated by a hyper-feature or a selected major node must have one and only one value.

There is a further point that needs to be made explicit about this view of feature relations. It is only if a segment has to be fully specified that one or more of the major nodes in each set has to be selected, and that each properly dominated feature has to have a value. As Keating (1988) has suggested, the output of the phonological component of a grammar may contain sounds that are completely unspecified for certain features. Thus a glottal stop or [h] may have no phonological specifications for the place node, although, of course, the full phonetic specification of [ʔ] and [h] must have some value for the place features; the tongue must be in *some* position for every sound. Similarly, although an epiglottal stop cannot be contrastively [+nasal] or [-nasal], nevertheless the soft palate must be either up or down during its production, so that in a full *phonetic* specification there is some value for the feature Nasal.

The set of possible phonological feature combinations is severely limited by the rules concerning possible paths through the hierarchical structure of hyper-features and major nodes. There are also very considerable constraints built in as a result of using multivalued features. Half of the 16 terminal features in Figure 3.1 are multivalued. With binary features all these distinctions can be realized only by using more features; but these additional features will always produce combinations of terms, such as *[+high, +low], which will have to be ruled out as definitionally impossible. Furthermore, as we noted in Chapter 2 and elsewhere, the introduction of unmotivated binary features eliminates the advantages of a linearly ordered set (such as [high], [mid], [low]), which enables us to write more explanatory rules.

There are, however, many additional limitations on the possible phonological specification of segments. Some of these are absolute constraints. For example, if a segment is [+voice] it must also have one of the values [creaky], [modal] or [breathy] for the feature Glottal Aperture; and if it is specified for Labial or Radical then it must be [-lateral]. Some other combinations of feature values are best regarded as phonological impossibilities. For example ejective nasals (to use a shorthand label for [stop, +nasal, ejective] segments) can be made, but they certainly do not appear.

We should also note that there are combinations of values of features that can be used as ways of distinguishing the sounds of one language from those of another, which have not been observed to be used contrastively within a single language. For example there is no known contrast between a voiceless alveolar lateral fricative [ɬ] and a voiceless alveolar lateral approximant [l̥]; but Maddieson and Emmorey (1984) have shown that some languages consistently use one of these possibilities and others the other. Distinctions such as these should be given some special status (or perhaps omitted altogether) in a theory providing an account of all possible phonologically contrastive segments.

Bearing all these points in mind, we must obviously regard Figure 3.1 as only a limited part of a theory specifying the set of possible phonological segments. What is needed in order to complete this part of linguistic theory is a set of rules showing, perhaps by co-indexing, those feature combinations that are impossible and those that are highly unlikely. Underlying this there should be a theory of phonetic naturalness, a formal statement of the old idea of ease of articulation combined with some measures of auditory distinctiveness. We are a long way from being able to say what is and what is not phonologically possible; but Figure 3.1 is a step towards organizing our knowledge.

Finally, we should consider the relationship between the auditory features discussed in chapter 2, which are intended simply to define some of the natural classes required in rules, and the articulatory features listed in this chapter which are also required to be sufficient to specify all the distinct lexical items in the world's languages, and to be necessary for constraining the set of possible phonological segments. The dependencies between the features in these two sets seem to be as shown in Figure 3.8. Some of them are fairly straightforward. Brightness and Height require a Dorsal specification; and Sibilant requires a Coronal specification. Sonorant acts as if it were a type of Stricture, and is accordingly made a sister of Aperture and Lateral. This leaves only Grave, which, in the system being defined here, is restricted to voiceless obstruents. This, like some of the other dependencies we have discussed, cannot be shown simply in the tree structure — some form of co-indexing statements will be required. For the moment Grave is just listed as a separate feature.

Adding all these features to the tree makes it a very redundant structure; but if we want to explain why languages are the way that they are, we have to take these auditory properties of sounds into account. Future research may show that some pruning of the tree is possible. But we must recognize that languages demand the groupings of some sounds in auditory terms, and that Sonorant, for example, is a necessary phonological feature despite the fact that it does not fit neatly into an articulatory feature tree. Languages have lived a long time, developed intricate phonological structures, and grown complex. We cannot always expect simplicity.

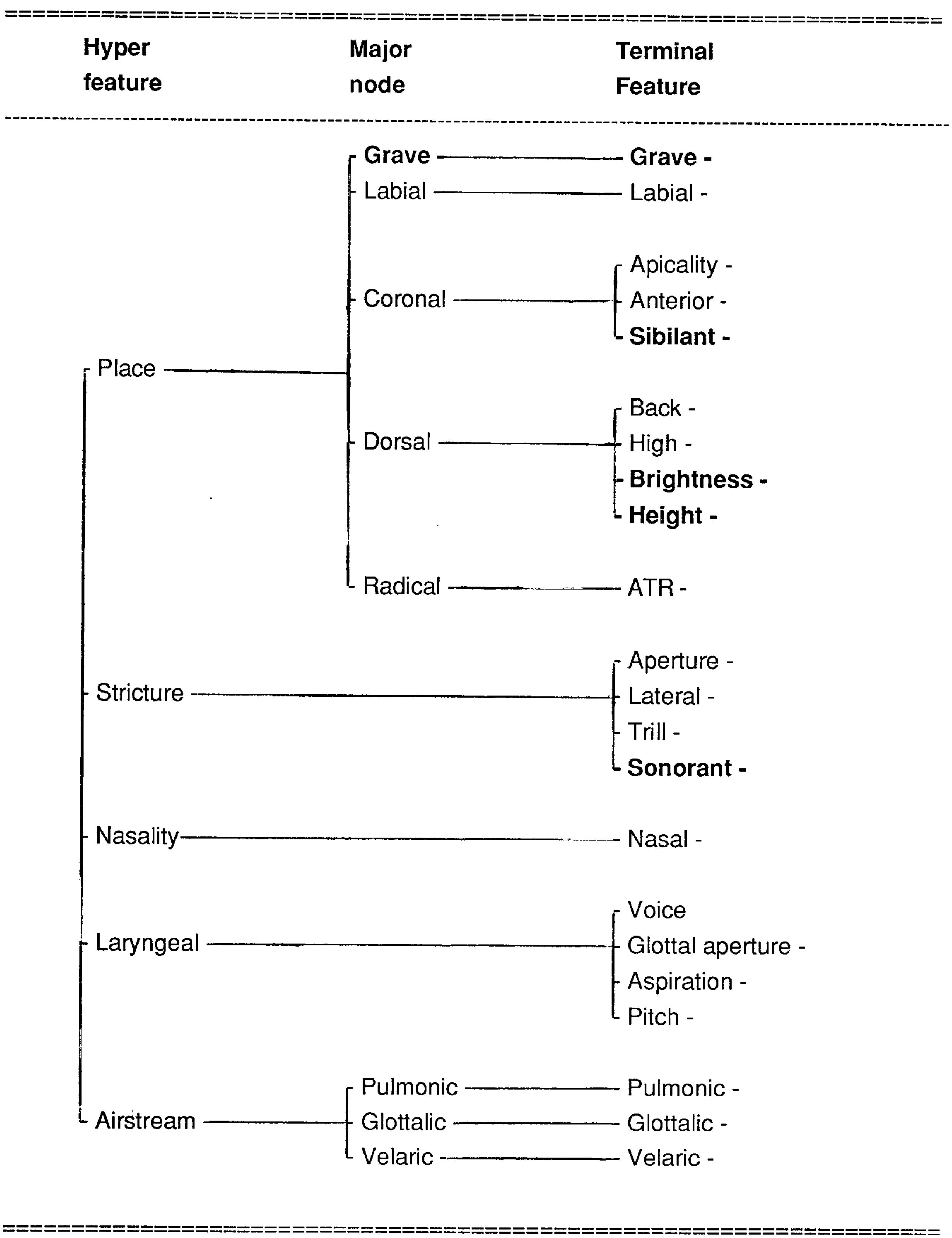

Figure 3.8 A feature tree, showing the relation between articulatory and auditory features. The auditory features are shown in bold face.

References

Abercrombie, David. (1967). Elements of general phonetics. Edinburgh: Edinburgh University Press.

Anderson, Steven R. (1978). Tone features. Tone: a Linguistic Survey. (ed. V. Fromkin), 133-176. New York: Academic Press.

Antilla, R. (1972). An Introduction to Historicial and Comparative Linguistics. New York: Macmillan.

Basböl, Hans. (1974). The phonological syllable with special reference to Danish. Annual Report of the Institute of Phonetics, Copenhagen, 8, 39-128.

Bell, A. Melville. (1867). Visible speech. London: Simpkin, Marshal.

Bladon, R. Anthony, and C. Gunnar Fant. (1978.) A two-formant model and the cardinal vowels. Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm 1978/1:1-8.

Blumstein, Sheila, E. and Kenneth N. Stevens. (1979). Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. Journal of the Acoustical Society of America. 66, 1001-1017.

Bolinger, Dwight. (1877). Another glance at main clause phenomena. Language. 53.3, 511-520.

Bolla, Kalman. (1980). A phonetic conspectus of Hungarian. The articulatory and acoustic features of Hungarian speech-sounds. Magyar Fonetikai Füzetek (Hungarian Papers in Phonetics). Vol.6.

Catford, John C. (1977a). Fundamental problems in phonetics. Bloomington and London: Indiana University Press.

Catford, John C. (1977b). Mountain of tongues: the languages of the Caucasus. Annual Review of Anthropology 6:283-314.

Chomsky, Noam. (1975). Reflections on Language. New York: Parthenon Press.

Chomsky, Noam and Morris Halle. (1968). Sound pattern of English. New York: Harper and Row.

Clements, G.Nicholas. (1985). The geometry of phonological features. Phonology Yearbook 2: 225-52.

Dart, Sarah. (1987). An aerodynamic study of Korean stop consonants: Measurements and modeling. Journal of the Acoustical Society of America 81.1:138-147.

Dauer, Rebecca M. (1983). Stress-timing and syllable-timing reanalyzed. Journal of Phonetics 11. 51-62.

Delattre, Pierre. (1965). Comparing the phonetic features of English, German, Spanish and French. Heidelberg:Julius Gross Verlag.

Disner, Sandra F. (1980). Evaluation of vowel normalization procedures. Journal of the Acoustical Society of America. 67.1, 253-261.

Dixit, R. Prakash. (1987a). Mechanisms for voicing ans aspiration: Hindi and other languages compared. UCLA Working Papers in Phonetics. 67, 49-102.

Dixit, R. Prakash. (1987a). In defense of the phonetic adequacy of the traditional term 'voiced aspirated'. UCLA Working Papers in Phonetics. 67, 49-102.

Dogil, Grzegorz. (1988). Linguistic phonetic features: a study in systematic phonetics. L. A. U. D: Duisburg.

Doke, Clement M. (1931). A comparative study in Shona phonetics. Johannesburg: University of the Witwatersrand Press.

Emeneau, Murray B. (1984). Toda grammar and texts. Philadelphia: American Philosophical Society.

Fischer-Jørgensen, Eli. (1952). Phonetic basis for identification of phonemic elements. Journal of the Acoustical Society. 24, 611.

Fischer-Jørgensen, Eli. (1985). Some basic vowel features, their articulatory correlates, and their explanatory power in phonology. Phonetic Linguistics, ed. by Victoria Fromkin. Orlando: Academic Press.

Galileo, Galilei (1633). [As reported in traditional accounts of his trial for heresy.]

Gershwin, Ira and George Gershwin. You say either, and I say either.

Gould, Stephen, J. (1981). The Panda's Thumb. New York: Norton.

Goldsmith, John. (1979). Autosegmental phonology. Garland: New York.

Halle, M. (1964). On the bases of phonology. In The Structure of Language, ed by Jerry A Fodor and Jerrold J. Katz. Englewood Cliffs, New Jersey: Prentice Hall. 324-333.

Halle, Morris. (1983). On distinctive features and their articulatory implementation. Natural Language and Linguistic Theory 1:1, 91-105.

Halle, Morris. (1988). The immanent form of phonemes. Giving birth to cognitive science: a festschrift for George A. Miller, ed. by William Hirsp. Cambidge, U.K.: Cambridge University Press.

Halle, Morris and Kenneth N. Stevens. (1971). A note on laryngeal features. Quarterly Progress Report of the Research Laboratory of Electronics, M.I.T. 101:198-213.

Halle, Morris and J-R. Vergnaud. (1980). Three dimensional phonology. Journal of Linguistic Research. 1.1, 83-105.

Hardman, Michael J. (1966). Jaqaru: Outline of phonological and morphological structure. (Janua Linguarum, Series Practica, 22.) The Hague: Mouton.

Harshman, Richard, Peter Ladefoged and Louis Goldstein. (1978). Factor analysis of tongue shapes. Journal of the Acoustical Society. 62, 693-707.

Holder, William. (1669). The elements of speech. London.

Hockett, Charles. (1955). A Manual of Phonology. Indiana University Publications in Anthropology and Linguistics.

Hulst, Harry van der. (1988). The phonetic and phonological basis of the simplex feature hypothesis. Sixth International Phonology meeting, Krems, Austria.

Hyman, Larry. (1975). Phonology theory and analysis. New York, Holt, Rinehart and Winston.

Ingram, John, C. (1975). Perceptual dimensions of phonemic recognition. Ph.D. dissertation, University of Alberta, Edmonton, Alberta.

Inouye, Susan Banner. (1988). The representation of internal segment structure. M.A. thesis, UCLA.

Jakobson, Roman, C. Gunnar Fant, and Morris Halle. (1951). Preliminaries to speech analysis. Cambridge: M.I.T. Press.

Jakobson, Roman, and Morris Halle. (1956). Fundamentals of language. The Hague: Mouton.

Jaeger, Jeri J. (1986). Concept formulation as a tool for linguistic research. Experimental phonology, ed. by John J. Ohala and Jeri Jaeger, 211-235. Orlando: Academic Press.

Joos, Martin. (1950). Description of language design. Journal of the Acoustical Society. 22, 701.

Keating, Patricia A. (1984). Phonetic and phonological representation of stop consonant voicing. Language 60:2, 286-319.

Keating, Patricia. (1988). Underspecification in phonetics. Phonology Yearbook 5:2.

Kim, Chin-Wu. (1970). A theory of aspiration. Phonetica. 21, 107-116.

Kirk, Paul, Peter Ladefoged and Jenny Ladefoged. (1984). Using a spectrograph for measures of phonation types in a natural language. UCLA Working Papers in Phonetics.59, 102-113.

Kisseberth, Charles. (1970). On the functional unity of phonological rules. Linguistic Inquiry. 1, 291-306.

Ladefoged, Peter. (1967). Three areas of experimental phonetics. London: Oxford University Press.

Ladefoged, Peter. (1968). A phonetic study of West African languages. (2nd ed.) Cambridge: Cambridge University Press.

Ladefoged, Peter. (1971). Preliminaries to linguistic phonetics. Chicago: University Press.

Ladefoged, Peter. (1975). A course in phonetics. New York: Harcourt Brace Jovanovich, Inc.

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

Ladefoged, Peter. (1982). A course in phonetics (2nd ed.). New York: Harcourt Brace Jovanovich, Inc.

Ladefoged, Peter and Peri Bhaskararao. (1983). Non-quantal aspects of consonant production: A study of retroflex consonants. Journal of Phonetics 11: 291-302.

Ladefoged, Peter, Anne Cochran and Sandra Disner. (1977). Laterals and trills. Journal of the International Phonetic Association 7.2, 46-54.

Ladefoged, Peter and Donald Broadbent. (1960). Perception of sequence in auditory events. Quarterly Journal of Experimental Psychology. 12, 162-170.

Ladefoged, Peter, and Ian Maddieson. (1986). The sounds of the world's languages. UCLA Working Papers in Phonetics 65.

Ladefoged, Peter, Ian Maddieson and Michel Jackson. (1987). nvestigating phonation types in different languages. " In Vocal Physiology: Voice Production, Mechanisms and Functions, ed. O Fujimura. Raven, New York. 297-317.

Ladefoged, Peter and Anthony Traill. (1980). Phonological features and phonetic details of Khoisan languages. Bushman and Hottentot Linguistic Studies (Papers of seminar held on 27 July 1979, ed. J.W. Snyman). Pretoria: University of South Africa. 132-166.

Ladefoged, Peter and Anthony Traill. (1984). Linguisic phonetic descriptions of clicks. Language 60:1, 1-20.

Ladefoged, Peter, Kay Williamson, Ben Elugbe, and A. Uwalaka. (1976). "The stops of Owerri Igbo. Studies in African Linguistics, Supp. 6, 147-163.

Ladefoged, Peter and Z. Wu. (1984). Places of articulation: an investigation of Pekingese fricatives and affricates. Journal of Phonetics 12, 267-278.

Lahiri, Aditi, Letitia Gewirth and Sheila E. Blumstein. (1984). A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: Evidence from a cross-language study. Journal of the Acoustical Society of America. 72.2, 391-404.

Laver, John. (1980). The Phonetic Description of Voice Quality. Cambridge, Cambridge University Press.

Lerner, Alan J. and J. Loewe. (1956). My Fair Lady. New York, Coward-McCann.

Liberman, Alvin M. and Ignatius G. Mattingly. (1985). The motor theory of speech perception revised. Cognition.

Lieberman, Philip. (1970). Towards a unified phonetic theory. Linguistic Inquiry. 1.3, 307-322.
Liljencrants, Johan, and Björn Lindblom. (1972). Numerical simulation of vowel quality systems: The role of perceptual contrast. Language 48:839-862.

Lindau, Mona. (1984). Phonetic differences in glottalic consonants. Journal of Phonetics. 12, 147-155.

Lindau, Mona. (1985). The story of /r/. Phonetic Linguistics. ed. Victoria Fromkin. Orlando, Academic Press

Lindau, Mona and Peter Ladefoged. (1986). Variability of Speech Processes. In Invariance and Variability of Speech Processes, ed. Joseph Perkell and Dennis Klatt. Hillsdale, N.J. Lawrence Erlbaum. 464-478.

Lindblom, Björn. (1984). Can the models of evolutionary biology be applied to Phonetic Problems? In Proceedings of the Tenth International Congress of Phonetic Sciences. ed. Marcel P.R Van den Broecke and Anthony Cohen. Dordrecht, Foris. 67-82.

Lindblom, Björn. (1986). Phonetic Universals in vowel systems. In Experimental Phonology. ed. John J. Ohala and Jeri J. Jaeger. Orlando, Florida: Academic Press, Inc. 13-42.

Linker, Wendy. (1982). Articulatory and acoustic correlates of labial activity in vowels: a cross-linguistic study. UCLA Working Papers in Phonetics. 56.

Maddieson, Ian. (1984). Patterns of sounds. Cambridge: Cambridge University Press.

Maddieson, Ian and Karen Emmorey. (1984.) Is there a valid distinction between voiceless lateral approximants and fricatives? Phonetics 41, 181-190.

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

Miller, George and Patricia Nicely. (1955). An analysis of perceptual confusions among English consonants. Journal of the Acoustical Society of America. 27: 338-352.

Mohanan, K.P. and Tara Mohanan. (1984). Lexical phonology of the consonant system in Malayalam. Linguistic Inquiry 15, 575-602.

Morais, J. et al. (1979). Does awareness of speech as a sequence of phones arise spontaneously? Cognition 7, 323-331.

Moskowitz , Breyne A. (1973). On the status of vowel shift in English. In T. E. Moore, ed., Cognitive development and the acquisition of language, 223-260. New York: Academic Press.

Nartey, Jonas. (1982) On fricative phones and phonemes. UCLA Working Papers in Phonetics. 55, 1-181.

Nihalani, Paroo. (1983). Voice quality and its implications. Abstracts of the tenth International Congress of Phonetics Sciences. 743. Dordrecht, Holland: Foris.

Ohala, John J. (1974). Experimental historical phonology. Historical linguistics II. Theory and description in phonology, ed. by J.M. Anderson and C. Jones., 353-389. Amsterdam: North Holland Publishing Co.

Ohala, John J. (1981). The listener as a source of sound change. chicago Linguistic Society. Papers from the parasession on language and behavior. 178-203.

Ohala, John J. and Jeri Jaeger. (1986). Experimental phonology. Orlando: Academic Press.

Pierrehumbert, Janet B. (1979). The perception of fundamental frequency declination. Journal of the Acoustical Society of America. 66, 363-369.

Pike, Kenneth L. (1943). Phonetics. Ann Arbor: University of Michigan Press.

Prince, Alan (1980). A metrical theory for Estonian quantity. Linguistic Inquiry. 3, 511-562.

Sagey, Elizabeth C. (1986). The representation of features and relations in non-linear phonology. Ph.D. dissertation, M.I.T., Cambridge, Mass.

Sapir, Edward. (1949). The psychological reality of phonemes. In: Selected Writings of Edward Sapir. (ed. D. G. Mandelbaum) Berkeley: University of California Press. 46-50.

Schane, Sanford. (1973). Generative Phonology. Englwood Cliffs, New Jersey: Prentice-Hall.

Shadle, Christine H. (1985). The acoustics of fricative consonants. Ph.D. thesis, M.I.T.

Shaw, George Bernard. (1920). Pygmalion: a romance in five acts. London, Constable.

Shepard, R.N. (1972). Psychological Representation of speech sounds. Human communication: A unified View. Ed. by Edward E. David and Peter B. Denes. New York: McGraw-Hill.

Singh, Sadanandh and J. W Black. (1966). Study of twenty six intervocalic consonants as spoken and recognized by four language groups. Journal of the Acoustical Society of America. 39, 639-656.

Snyman, Jan W. (1975). Zul'hõasi fonologie en woordeboek. Cape Town: Balkema.

Stalin, Josef V. (1950). Marxism and problems of linguistics. Pravda June 20, 1950. Moscow.

Stevens, Kenneth N. (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In P.B. Denes and E.E. David Jr., eds., Human Communication, A Unified View. New York: McGraw-Hill.

Stevens, Kenneth N. (1983). Design features of speech sound systems. In The Production of Speech. Ed. Peter MacNeilage. New York, Springer Verlag. 247-262.

Stevens, Kenneth N. (1988). Modes of vocal-fold vibration based on a two-section model. Vocal physiology: Voice Production, Mechanisms and Functions, ed. by Osamu Fujimura. New York: Raven Press.

Stevens, Kenneth N. and Sheila E. Blumstein. (1975). Quantal aspects of consonant production and perception: a study of retroflex consonants. Journal of Phonetics 3, 215-233.

Stevens, Kenneth N. and Sheila E. Blumstein. (1981). The search for invariant acoustic correlates of phonetic features. in Perspectives on the study of speech. eds. P. D. Eimas and Joanne L. Miller. Hillsdale, New Jersey, Lawrence Erlbaum Associates. 1 -38.

Stevens, Kenneth N. S., Jay Keyser and Haneko Kawasaki. (1986). Toward a phonetic and phonological theory of redundant features. In Joseph S. Perkell and Dennis H. Klatt, eds., Invariance and variability in speech processes. Hillsdale, NJ: Lawrence Erlbaum Associates. Pp. 426-449.

Stockwell, Robert P. (1960) The place of intonation in a generative grammer of English. Language 36, 360-367.

Sweet, Henry. (1877). A handbook of phonetics, including a popular exposition of the principles of spelling reform. Oxford: Clarendon Press.

Terbeck, Dale. (1977). A cross-linguistic multi-dimensional scaling study of vowel perception. UCLA Working Papers in Phonetics. 37.

Trager, George and Henry Lee Smith. (1951). An Outline of English Structure. Norman, Oklahoma.

Traill, Anthony. (1985). Phonetic and phonological studies of !Xóõ Bushman (Quellen zur Khoisan-Forschung 1). Hamburg, Helmut Buske.

Trubetzkoy, N.S. (1929). Zur allgemeinen theorie der phonologischen vokalsysteme. Travaux du Cercle Linguistique de Prague 1:39-67.

Trubetzkoy, N.S. (1939). Grundzüge der Phonologie. Travaux du Cercle Linguistique de Prague 7.

Wittgenstein, Ludwig. (1958). The Blue and Brown Books. Preliminary studies for the Philosophical Investigations. New York: Harper and Row.