

The Sounds of the World's Languages

The title of this book, *The Sounds of the World's Languages*, implies two very significant claims. One is that it makes sense to talk about entities that can be labeled 'sounds'. The other is that we know enough about the languages of the world to be able to write a book that covers them all. We would like to explain our thinking in selecting this title.

Most of the phonetic literature of the last two centuries assumes that a meaningful analysis can be made by dividing speech into small chunks that can be called speech sounds. These may be described in formalized terms as phonemes, root nodes, or some other theoretical entity, or discussed in more general terms as 'segments'. The tradition of segmental description has proven its utility in many ways. Nonetheless, linguists' views on what segments are have varied greatly. At one extreme are those who regard them as descriptive fictions, invented solely by the linguistic analyst. At an opposite extreme are those who hypothesize that the organization of speech production, and the structure of words in memory are based directly on segments. We do not feel it necessary in this book to take a stand on either side of this issue, but choose to employ segmental descriptions of languages as the most suitable way of communicating the results of phonetic analyses to a wide range of readers.

The second part of the title, the world's languages, also deserves some comment. We believe that enough is now known to attempt a description of the sounds in all the languages of the world. Professional linguists have some knowledge of the great majority of the languages now spoken. This is a state of

affairs which has never existed at any time prior to the latter part of this century. The 'global village' effect means that few societies remain outside the scope of scholarly scrutiny. In all probability there will be a sharp decrease in the rate at which previously unknown sounds are drawn to the attention of phoneticians. We are, of course, aware that there are phonetic phenomena in every language that have yet to be described. Speech varies in response to many different circumstances, and we do not have a complete knowledge of the phonetic structure of any language. In addition, languages are always evolving. Thus there can never be a final description of the sounds of any one language. The next generation of speakers will always speak a little differently from their predecessors, and may even create sounds that have never been used in a human language before. We think it probable, however, that any new sounds will be similar to those that now have a linguistic function and will be formed by re-arrangements of properties of sounds that have been previously observed in linguistic usage. In other words, we feel that a basis exists for discriminating between linguistic and non-linguistic sounds.

Using this basis and our intuitions we have sometimes posited the existence of sounds that have not yet been reported in the linguistic literature. These are sounds which we feel reflect accidental gaps in the currently available data, or are absent only by chance from any currently spoken language. Other possibilities are not mentioned at all since we believe they will never have a role in linguistic structure. There are, of course, many sounds that can be made with the vocal organs that are not known to be used in any language. People can whistle, click their teeth, wag their tongues from side to side, and perform a variety of other maneuvers to produce sounds that have never been reported to have a linguistic function. But linguistic phonetics does not have to account for all the sounds that humans are capable of making, or even all of those which can be made just in the vocal tract.

The primary data we will try to describe are all the segments that are known to distinguish lexical items within a language. We have in this way determined the level of description at which we will operate. We are concerned with the lexical segments that account for minimal pairs. (We must admit that in a number of cases we have insufficient knowledge of the phonology of the language being described to be absolutely sure of the nature of all apparent contrasts; but in general, information at the level of minimal pairs is relatively accessible to both the native speaker and the linguist.) The segments we have chosen to describe may not be the same as the underlying phonological segments. Thus standard varieties of English have minimal sets such as *rum*, *run*, *rung*, and hence a contrastive inventory that includes three nasal segments, **m**, **n**, **ŋ**. The third of these, the velar nasal **ŋ**, has a different distributional pattern from the other two. It never appears initially in a morpheme and it alternates with **ng** before the comparative adjectival suffix *-er* in words like *long*, *longer*. Words like *dim*, *dimmer*; *thin*, *thinner* do not show a parallel alternation between nasal and nasal plus stop. For several reasons of this sort the segment **ŋ**

can be said to be absent from the underlying inventory of English, although it serves to establish minimal contrasts. The different status of η is not of concern to phonetics. We want to account for all the contrasts that occur between distinct segments, and in English that includes η .

There are, of course, other situations which are harder to resolve in this way. In particular it is often difficult to decide how to divide a particular phonetic string into segments. There are well-known cases such as the affricates $tʃ$ and $dʒ$ in English, which some linguists regard as one segment and others as a sequence of two segments. From our point of view it does not really matter which of these solutions is chosen. The phonetic facts to be described remain the same in either case. English also provides a useful illustration of another form of this problem in words such as *mew* and *beauty*. Obviously these words differ from *moo* and *booty*, but the difference could be interpreted as a contrast between the presence and absence of the semivowel j , or as a contrast between the different vowel nuclei $i\bar{u}$ and u (where $i\bar{u}$ represents a diphthong with a rising amplitude contour). However, again the phonetic facts would not differ according to which view is taken. Our approach in similar situations is to adopt the phonetic description which yields the simplest segment types. Thus, in this case, we would prefer the interpretation with the simpler segment j to the complex one $i\bar{u}$.

Although our primary data are the contrasts within languages, we also take note of differences between languages. We hope that the phonetic events observed will be sufficient to form the basis for an overall phonetic theory. This needs to be rich enough to describe those segmental events which distinguish one language or accent from another and which are also sufficiently distinct to serve as potential conveyers of lexical contrasts for speakers of other languages. We will, however, mainly restrict the discussion to those segments that distinguish words in some language.

As we will be discussing only segments, prosodic features such as tone, stress and accent fall outside the scope of the book, even though they distinguish words. So also do most of the variations in segments that result from differences in their prosodic context. Furthermore, variations due to stylistic choices and changes in speech tempo are disregarded. We are implicitly restricting our attention to a careful speech style, and many of our examples are drawn from citation forms, or words spoken in carrier phrases. For the task we have in mind this choice is appropriate as these forms are likely to be the most differentiated at the segmental level.

Our idea of what a phonetic description of these events should be like involves the establishment of parameters along which variation can be measured and a set of categorial values along these parameters. The categorial labels provide a vocabulary that can be used to classify distinctions within a language and to describe cases where sounds in different languages are sufficiently similar to be equated with one another. The parameters and categories that we use are mostly quite familiar ones from the phonetic literature, but we

have in each case tried to rethink, through our examination of contrasts in languages, the basis on which any given one is justified. We also discuss the relationship that exists between the phonetic categories we establish and the classificatory features employed in phonological theory. In these discussions we take as our usual point of reference the features that have been used in recent models of feature geometry (see the review by McCarthy 1988), although we also pay some attention to some older proposals and certain newer suggestions. We do not attempt to present our own competing phonological feature system, as we are not examining phonological processes and constraints, but only contrasts (and even these from a largely phonetic point of view). We do, however, point to some areas where we believe that some elaboration or clarification of the usual feature scheme might be advisable. In the Coda at the end of the book we summarize the major findings on segmental contrasts which a phonological theory must address.

We also maintain that a distinction can be drawn between those phonetic parameters that are likely to serve as a basis for the organization of linguistic behavior, and others that can be used to provide a physical characterization of sounds, but which do not relate well to the sound patterns of languages. We could, for example, consider place of articulation to be defined in terms of a single continuous parameter such as the distance of the stricture from the glottis. However, this description would overlook, among other things, the functional and anatomical separation that exists between the lips and the tongue. A more insightful description is therefore one that partitions place of articulation into a number of major zones, within each of which certain categories can be distinguished.

As is detailed in our discussion of place of articulation in chapter 2, this does not mean that boundaries between each category can be easily determined. For example, the tongue has no clearly defined regions, and neither does the roof of the mouth; and there are, of course, similar problems in dividing the continua which underlie other phonetic parameters. Nonetheless, it is very striking that languages often seem to cut the continua in similar ways so that it is possible to equate a sound in one language with a similar sound in another. Our procedure in trying to build up a description of the sounds of the world's languages is to proceed primarily by a series of such equivalences.

Thus, we might encounter two languages A and B, each of which has a set of three sounds that differ only by differences along a single parameter. In each language it will be possible to rank the three sounds according to their values on the parameter in question, A1, A2, A3 and B1, B2, B3. Now, if A1 is judged to be equivalent to B2, and A2 equivalent to B3, there must be (at least) four values of the parameter in question which suffice to distinguish contrastive sounds. The equivalences can be displayed as follows:

1			B1
2	A1	=	B2
3	A2	=	B3
4	A3		

The likely total of contrastive categories along a given parameter is established by considering a set of such equivalences. Of course, judging which two sounds in different languages are equivalent rests on both the evaluation of data and experience.

A particular speech sound is not characterized by a value on only one parameter, but by the set of specified values on all the relevant parameters. We organize our discussion primarily on the basis of integral sounds, not according to single parameters or groups of related parameters. For example, there are chapters devoted to stops, nasals, fricatives, and vowels, rather than chapters which discuss topics such as degrees of stricture, types of phonation, and lip position. There is one significant exception to this general practice: In chapter 2 we survey the places of articulation needed to characterize the sounds of the world's languages, focussing on place contrasts found in stops and nasals. This chapter provides a framework for description of place contrasts among other classes of segments.

The next chapter, chapter 3, describes the remaining aspects of the various types of stops that occur (apart from clicks, to which a separate chapter is devoted). This chapter contains the principal discussion of distinctions in laryngeal activity between segments. The four succeeding chapters discuss other major types of consonants – nasals, fricatives, laterals and rhotics – in terms of the distinctions in place and laryngeal activity previously established, as well as any additional parameters that relate to the specific class of consonants covered by the chapter. Chapter 8 is devoted to clicks. Although sounds of this type occur in relatively few of the world's languages, there is great variety among clicks due to the fact that a large number of ways of modifying them are employed in the languages that use them. The survey then continues with a chapter on the vowels and semivowels observed in the world's languages. A final chapter, on multiple articulations, discusses single segments produced by combining several articulations each of which on its own could have formed a segment. This chapter includes discussion of what are traditionally called secondary articulations.

In these chapters we treat as separate categories only those differences between sounds that occur within the same language and those differences between sounds in different languages that we believe are great enough to be potentially distinctive. However, a descriptive system of phonetic parameters should also be able to characterize many differences at the phonetic level that are unrelated to phonological contrast. Differences in regional or social accents can be maintained by, for example, subtle shifts in vowel quality that are noticeable over a long stretch of speech, but which are not sufficient for

distinguishing the meanings of words or phrases. That is, a phonetic theory should be able to make a distinction between those properties of sounds that can or could convey linguistic information (the difference between one form and another within a possible language) and those that convey only sociolinguistic information (the difference between one language or dialect and another). We believe that it is possible to set up a system of phonetic parameters each of which contains categories that taken as a whole, will distinguish all the potential contrasts within human languages. The system also allows for differences between one language and another to be described by different values along the parameters.

In addition to the categories and parameters of our descriptions we sometimes refer to targets or gestures in our accounts of segments. We would like to make clear that we are using these terms without implying any specific theoretical framework. A target can be thought of as either a specific location towards which an articulator is moving or as a more general notion including several properties such as the rate of the movement or its acoustic results. In either case, a target is an abstract goal at which one aims but which one does not necessarily hit. A gesture in our usage is a movement of a single articulator, which may by itself form the basis of a segment, or may require coordination with other articulatory movements for the formation of the sound. This is not the same concept as used in Articulatory Phonology (Browman and Goldstein 1992) but is based on similar insights. We adopt this usage to provide a framework within which to discuss dynamic aspects of speech. Speech is not a static process, but an active one, and it is clear that many properties cannot be understood unless we examine their dynamic aspects.

Although the discussion in this introductory chapter and the general layout of this book are in terms of articulatory properties of sounds, throughout the book we have also taken note of acoustic properties. It is perfectly possible to describe all the systematic phonetic contrasts that occur among languages in terms of the sound-producing mechanism. But this is often not the appropriate way of characterizing contrasts; some phonological statements depend more on the acoustic properties of sounds. In any feature theory that is eventually developed, some features will classify sounds in terms of their acoustic properties, which are presumed to be in a simple relation with the auditory parameters that the perceptual mechanism uses. There is also growing evidence that the mental targets of some sounds (notably vowels) are best described in terms of an auditory theory of speech production (Johnson, Ladefoged and Lindau 1993).

There are many other points that should be taken into account in a theory of linguistic phonetics. We should note that some places of articulation (or some areas within the continuum) are used more often than others. Similarly certain manners of articulation (such as stops) are common, but others (such as trills) occur less often; and within the vowel continuum, some vowels are found in a

far wider range of languages than others. We do not at the moment have a complete theory that will account for these facts. There is no doubt that some articulations are easier to make than others. Thus although it is perfectly possible for anybody to make a closure between the blade of the tongue and the upper lip, this gesture is not used in many languages, perhaps because it is harder to integrate into the stream of speech than, say, a bilabial stop. There is also no doubt that some auditory distinctions are easier to maintain than others; the set of vowels *i, e, a, o, u* are more distinct than the set *y, ø, ə, ɤ, ʊ*. Furthermore, as Stevens (1989) has pointed out, some articulatory gestures can vary over a fairly wide range without producing much acoustic change, resulting in what he calls quantal articulations. But while pursuing these topics is undoubtedly of great importance for an adequate theory of phonetics, our aim in this book is more limited, and we will comment on them only in passing.

We have attempted to provide a great deal of documentation of the contrasts we discuss. As a result, we believe that a part of the value of the book lies in its exemplification of a wide variety of types of phonetic data collected by a range of techniques. The data includes still and cine x-ray photography, palatography, spectrograms, waveforms, aerodynamic data, and articulatory movement data obtained by a number of different methods. We have not attempted to explain how to use these methods, as this book is not a handbook of phonetics, but we hope that we have provided sufficient pointers so that readers will understand the significant aspects of the documentation provided. Readers are presumed to be familiar with basic phonetic concepts as expounded in standard textbooks. We are, however, aware that we refer to so many languages that readers are unlikely to be familiar with all of them. With this in mind, we provide an appendix identifying the location and classification of all languages mentioned. In the text itself languages will usually be cited without further identification.

As with any jointly authored work, this book represents the results of some compromises between the authors. Had only one of us been writing the book, some data would have been interpreted differently, and certain arguments would undoubtedly have been stated more forcefully than they appear here, while others would have been omitted altogether. But we both believe that the combination of our efforts has produced a more complete survey than either of us could have produced alone. When we started this project we knew that we were being ambitious. But we did not realize the extraordinary amount of physiological and acoustic data that is available on the little known languages of the world. Our library research has led us to hundreds of books and papers that contain wonderful instrumental data that has never been summarized or otherwise made available to the general phonetician who is not concerned with the particular languages being described in the original work. We also found that our own files and the UCLA phonetic archives accumulated over many years contained a great deal of previously unpublished material. We were often pleased to find that illustrative recordings and analyses of

particular sounds cited could be obtained from our own resources. The great wealth of material available to us led to some problems in trying to decide which of many pieces of data we should include. We wanted to write a book that would advance linguistic and phonetic theory. Accordingly we have limited ourselves to discussing just the data required for this purpose. We have not included a number of things that are well known and readily available, such as acoustic data illustrating the contrasting stop consonants in English, or extended discussions of voice onset time.

As our research went on, we also uncovered some notable gaps in the available published data and our own resources. This led to an interesting problem in research management. To what extent should we try to fill these gaps by extending our own investigations? Our solution to this problem was to set a date for the completion of this book, and do what we could within this time limit. As a result we are woefully aware that much more could have been done. We have been able to provide a large number of new analyses of our own previously unpublished data, and to summarize much of the literature. But we can still foresee a lifetime of work ahead of us.

This is the present state of our knowledge about the linguistic phonetic events that occur in the languages of the world.