OXFORD
LINGUISTICS

A THEORY OF PHONOLOGICAL FEATURES

SAN DUANMU

clicks
vowel
affricates
consonant
inventory database
features
ejectives
allophone
complex sounds
phoneme
vowel height

A Theory of Phonological Features

A Theory of Phonological Features

SAN DUANMU

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First Edition published in 2016

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Control Number: 2015948898

ISBN 978–0–19–966496–2

Printed in Great Britain by
Clays Ltd, St Ives plc

Contents

Preface

The goal of this study is to determine a feature system that is minimally sufficient to distinguish all consonants and vowels in the world's languages. Evidence is drawn from two databases of transcribed sound inventories, UPSID (451 inventories) and P-base (628 inventories).

Feature systems have been offered before that use data from many languages. For example, Trubetzkoy (1939) cited some 200 languages, Jakobson et al. (1952) cited nearly 70 languages, and Maddieson (1984) used a database of 318 languages. However, a fundamental methodological question remains to be addressed: Let X be a sound from one language and Y a sound from another language. How do we decide whether X and Y should be treated as the same sound or different sounds? It is well known that, when X and Y are transcribed with the same phonetic symbol, there is no guarantee that they are phonetically the same. Similarly, when X and Y are transcribed with different phonetic symbols, there is no guarantee that they must be different sounds. Moreover, even if there is phonetic evidence that X and Y are notably different, there is no guarantee that they cannot be treated as the same sound in a language. Without a proper answer to the methodological question, generalizations from cross-language comparisons are open to challenges.

In this study, I offer an answer, using the notion of contrast: X and Y are treated as different sounds if and only if they contrast in some language (i.e. distinguish words in that language). For example, [l] and [r] contrast in English (as in *lice* vs. *rice*); therefore, we must represent them with different transcriptions and different features. In addition, when [l] and [r] occur in a language where they do not distinguish words, such as Japanese, we can distinguish them, too (as "allophones"). On the other hand, if X and Y never contrast in any language, there is no need to distinguish them, not even as allophones, even if they have an observable phonetic difference. For example, Ladefoged (1992) observes a difference between [θ] used by English speakers in California, whose tongue tip is visible, and [θ] used by those in southern England, whose tongue tip is not visible; but if the two kinds of [θ] never contrast in any language, there is no need to distinguish them, either in transcription or in features. Similarly, Disner (1983) observes that the [i] in German has a slightly higher tongue position than that in Norwegian. But again, if the two kinds of [i] never contrast in any language, there is no need to distinguish them either. Non-contrastive differences are not left aside but will be addressed as well, and explanations will be suggested without assuming feature differences.

I make no prior assumption as to what the system should look like, such as whether features should be binary or innate. Instead, the proposed method is explicit and theory-neutral. The work can be laborious to carry out, though: It requires repeated searches through sound inventories of all available languages in order to find out whether a difference in phonetics or in transcription is ever contrastive in any language. In addition, to guard against errors in original sources or in the databases, we shall re-examine all languages that seem to exhibit an unusual contrast.

The resulting feature system is surprisingly simple: Fewer features are needed than previously proposed, and for each feature, a two-way contrast is sufficient. For example, it is found that, if we exclude other factors, such as vowel length and tongue root movement, a two-way contrast in the backness of the tongue is sufficient; the same is true for the height of the tongue. This result is quite unexpected, because even the most parsimonious feature theories, such as Jakobson et al. (1952) and Chomsky and Halle (1968), assume three degrees of tongue height, and many assume three degrees of tongue backness as well. Nevertheless, our conclusion is reliable, in that the notion of contrast is uncontroversial, the proposed procedure is explicit, and the result is repeatable.

Representing contrast is not the only purpose for which features are proposed. In particular, features have been proposed to describe how sounds are made (articulatory features) and how sounds form classes in the phonology of a language (class-based features) as well. This study focuses on contrast-based features for two reasons. First, contrast lies at the core of phonology. Second, contrast-based data are the least controversial and are large in quantity and high in quality. I shall attempt to interpret contrast-based features as articulatory gestures, too, but I shall say little of class-based features. It has been proposed that contrast-based features, articulatory features, and class-based features should correspond to each other (Halle 1995), but that remains an ideal. Before the ideal is confirmed, differences among the three feature systems, once highlighted, provide fertile grounds for further research.

Acknowledgments

The theory of phonological features is one of the first topics a student of linguistics comes across, yet one soon realizes that there are many unresolved questions. Naturally, many phonologists have searched for answers, and some have devoted most of their careers to it. If some progress is made in this book, it is very much the result of collective effort. I am glad to be part of it, and have many people to thank.

Starting from the beginning, I would like to thank those whose classes introduced me to the field of phonetics and phonology: David Crystal, Francois Dell, William Hardcastle, Jim Harris, Michael Kenstowicz, Dennis Klatt, Donca Steriade, Ken Stevens, Peter Trudgill, Moira Yip, Victor Zue, and above all Morris Halle. I was fortunate to have you as my teachers.

I would also like to thank my colleagues at the University of Michigan, in Linguistics and beyond. It is a privilege to share a great research and teaching environment with you.

Next, I would like to thank other phonologists, broadly defined. Some I have interacted with directly and some I got to know through their works. You have been wonderful companions in a journey of exploration for a better understanding of the sound patterns of language.

Just as much, I would like to thank my students over the years, whose open-minded questions often helped me clarify my explanations, revise my views, and see new questions that had been hidden by familiar jargons.

Finally, I would like to thank Michael Kenstowicz, Keren Rice, and three anonymous reviewers for their valuable comments on the book proposal, Bert Vaux and Jeff Mielke for their very careful comments on the draft copies of the book, other colleagues for their comments on various parts of the book (Steve Abney, George Allen, Hans Basbøll, Pam Beddor, Iris Berent, Arthur Brakel, Charles Cairns, Nick Clements, Andries Coetzee, Jennifer Cole, Stuart Davis, Ik-sang Eom, Shengli Feng, Mark Hale, Bruce Hayes, Jeff Heath, Yuchau Hsiao, Fang Hu, José Hualde, Di Jiang, Ping Jiang, Pat Keating, Hyoyoung Kim, Andrew Larson, Qiuwu Ma, Bruce Mannheim, James Myer, Wuyun Pan, Eric Raimy, Rebecca Sestili, Jason Shaw, Yaoyun Shi, Ken Stevens, Will Styler, Sally Thomason, Jindra Toman, Hongjun Wang, Hongming Zhang, Jie Zhang, Jun Zhang, Jianing Zhu, and Xiaonong Zhu—and this surely is an incomplete list, I am afraid), and the staff at (or working for) Oxford University Press, in particular Sarah Barrett, John Davey, Lisa Eaton, Julia Steer, Vicki Sunter, and Andrew Woodard, for their professionalism at every stage of the publication process. You have helped make this book far better than what was originally conceived.

For copyright permissions I am grateful to the following:

Cambridge University Press for Fig. 1.2 (Karl Zimmer and Orhan Orgun, "Turkish," *Journal of the International Phonetic Association* 22(1–2) (1992), p. 44), also reproduced as Fig. 4.1.

Sandra Disner for Fig. 1.3 (Sandra Ferrari Disner, "Vowel quality: the relation between universal and language-specific factors," doctoral dissertation, UCLA, 1983; distributed as UCLA Working Papers in Phonetics 58, p. 67, Fig. 3.24).

AIP Publishing LLC for Fig. 2.2 (Gordon E. Peterson and Harold L. Barney. 1952. "Control methods used in a study of the vowels," *Journal of the Acoustical Society of America* 24(2) (1952), p. 182, Fig. 8).

Abbreviations and terms used

Commonly used abbreviations

*	ill-formed
[ ]	a feature, often one with binary values, such as [round]; a phonetic transcription
ATR	advanced tongue root
C	consonant
CC	unit of two consonants; geminate consonant
CG	unit made of a consonant and a glide, such as [kw] or [k^w]
CL	unit made of a consonant and a liquid ([l] or [r]), such as [pl]
CV tier	tier representing consonant and vowel positions
G	glide
H	high tone, which is the same as the feature value [+H]
HL	falling tone, which is a combination of H and L
IPA	International Phonetic Association; International Phonetic Alphabet
L	low tone, which is the same as the feature value [–H]
LH	rising tone, which is a combination of L and H
N	nasal consonant
NC	unit made of N and C; pre-nasalized consonant
P-base	database of phoneme inventories and sound classes (Mielke 2004–7)
UPSID	UCLA Phonology Inventory Database (Maddieson and Precoda 1990)
V	vowel
VX	rime made of a vowel and another sound, such as VV [ai] or VC [an]
X	timing slot; unit of time
X tier	tier representing time units

Active Articulators

Articulator	Common name	Other terms
Lips	Lips (or lower lip)	Labial; Lower lip
Tip	Tip of the tongue	Coronal; Tongue blade
Body	Body of the tongue	Dorsal
Velum	Velum	Soft palate
Root	Root of the tongue	Radical
Glottis	Glottis	Vocal cords; Vocal folds; Laryngeal
Larynx	Larynx	

1

Introduction

A fundamental assumption in phonology is that, at some level of abstraction, speech is made of a sequence of sounds, or consonants and vowels. Under this assumption, numerous studies of a language begin by listing its consonants and vowels. The goal of this study is to determine a minimally sufficient feature system that can distinguish all consonants and vowels in the world's languages. Before I introduce the methodology in Chapter 2, it is necessary to address some preliminary questions: What are speech sounds? What are features? Can we compare sounds and features across languages? Do we have adequate data for the task?

1.1 Sounds and time

The discovery that words can be decomposed into sounds has made it possible to create writing systems in many languages, which in turn has had a profound impact on human civilization. Indeed, according to Goldsmith (2011: 28), phonemic analysis (the technique for figuring out the consonants and vowels of a language) remains the greatest achievement in phonology. Nevertheless, a number of questions remain.

1.1.1 Segmentation of speech

When we segment speech into consonants and vowels, we encounter two problems. First, there are prosodic properties, such as tones, that seem to be independent of, or attached to entities larger than, consonants and vowels. Second, phonetically, the boundaries between sounds are not always clear, because properties of one sound often spill into another (and vice versa)—a process called "coarticulation" in phonetics and "feature spreading" in phonology. For example, in *pan*, nasalization (a property of [n]) starts during [æ], not after it. Similarly, the tone of a syllable can extend to another, or shift from one to another. In response, Goldsmith (1976) proposes that speech is made of multiple tiers of features, each tier being autonomous. Features on the same tier have their own temporal sequence, but there is no common temporal sequence across all layers. The conclusion is that speech cannot be segmented into consonants and vowels. A similar view is expressed by Firth (1957)

A Theory of Phonological Features. First edition. San Duanmu.

and more recently by Fowler (2012), who suggests that we should focus on individual articulatory gestures instead of sounds or segments.

However, some well-known facts will be hard to account for if speech is not made of a sequence of sounds (after we set aside prosodic properties such as tone and syllable structure). For example, given two sounds A and B, languages can make a contrast between AB and BA, such as *tax* vs. *task*, *cats* vs. *cast*, *tea* vs. *eat*, and *map* vs. *amp*. In addition, it is possible to spell or transcribe speech with a sequence of letters or phonetic symbols, regardless of the language—a fact that would be quite unusual if speech is not made of a sequence of sounds. Moreover, language games can manipulate (i.e. move, insert, delete, or change) sound-sized units, and such games are found not only in languages that are written alphabetically, such as Pig Latin in English, but also in those that are not, such as Chinese (e.g. Chao 1931; Yip 1982; Bao 1990b). Finally, there is no evidence that it is easier to account for feature spreading (or coarticulation) if we reject consonants and vowels. Indeed, even within the multi-tiered approach to phonology, a special tier has been proposed that corresponds to traditional notions of consonants and vowels, such as the CV tier of McCarthy (1979) and Clements and Keyser (1983), or the X tier of Pulleyblank (1983) and Levin (1985). Therefore, I shall continue to assume that speech is, at some level of abstraction, made of a sequence of sounds.

1.1.2 *Granularity of segmentation*

A second problem in decomposing words into sounds is the granularity of segmentation: there is no agreement on how large (or small) a sound should be. For example, is the affricate [ts] one sound or two? Is the diphthong [ai] one sound or two? Is the triphthong [uai] one sound or three? Is [ʘx] (in the African language !Xóõ) one sound or two? Should the decisions be made on a language-specific basis? For example, can [ai] be one sound in some languages and two in others, even if it is phonetically the same?

Chao (1934) argues that phonemic analysis is inherently ambiguous and a unique solution is rarely possibly. The ambiguity has made some linguists doubt the reality of consonants and vowels. For example, after years of working on consonants and vowels, the prominent linguist Ladefoged (2001: 170) remarks that they are probably "scientific imaginations" after all.

Nevertheless, many ambiguities are resolvable if additional evidence is taken into consideration. For example, consider the syllable onset in Standard Chinese. If we exclude glides, only the following items are found [p p^h t t^h k k^h ts ts^h tʂ tʂ^h tç tç^h m n f s ʂ ç x ɹ]. If the affricates [ts ts^h tʂ tʂ^h tç tç^h] are single consonants, we see a generalization: The Chinese onset allows just one consonant. If affricates are clusters of two sounds each, the generalization is lost. In addition, we face a new question whose answer is not so obvious: Why are some consonant clusters allowed while others not?

However, additional evidence is often ignored. As an example, consider two analyses of Chinese. You et al. (1980) propose that Chinese should not be segmented into consonants and vowels. Instead, we can treat each rime as a single sound, such as [au], [ai], [an], and [aŋ]. The advantage, they argue, is that we do not need to account for contextual variations of vowels (called "allophonic variation" in phonemic analysis), such as the variation of [a] in [au], [ai], [an], and [aŋ]. However, they fail to account for the fact that diphthongs and VC rimes are found in full syllables only, whereas rimes of unstressed syllables are half as long and have a simple vowel only (Lin and Yan 1988). If the difference between full and unstressed syllables is taken into consideration, it is clear that the former have two rime positions and the latter just one, so that [au], [ai], [an], and [aŋ] should all be split into two sounds each.

Next we consider vowels in Standard Chinese. According to Lee and Zee (2003), there are six monophthongs (such as [a] in [ma] 'mother' and [an] 'peace'), eleven diphthongs (such as [ai] in [mai] 'sell' and [ia] in [ia] 'duck' and [ian] 'smoke'), and four triphthongs (such as [uai] in [xuai] 'bad'). However, the analysis becomes problematic when we consider evidence from syllable structure. First, [mai] 'sell' and [xuai] 'bad' form a riming pair, as do [an] 'peace' and [ian] 'smoke'. This means that [uai] should be divided into [u] and [ai], because [ai] is a unit for riming. Similarly, [ian] should be divided into [i] and [an], where [an] is a unit for riming. Second, diphthongs like [ai] and [au] cannot be followed by a consonant, such as *[ain] or *[aun], whereas simple vowels can, such as [in] and [an]. This means that a diphthong is equal to two sounds. Thus, evidence from syllable structure suggests that both diphthongs and triphthongs should be decomposed into simple vowels. The decomposition yields better phonemic economy, too. For example, without decomposing diphthongs and triphthongs, there are twenty-one vowels in Standard Chinese (Lee and Zee 2003), whereas with the decomposition there are at most six, including a retroflex vowel (Duanmu 2007: 41).

Even if additional evidence is used, its interpretation is not always obvious. Let us consider two examples. We have just seen that diphthongs can be treated as two sounds when they are long. However, Cairns (p.c. 2013) suggests that some diphthongs are short. The example is New Yorkers' pronunciation of the vowel in *bath* and *cab* as [æᵊ]. Cairns considers the vowel to be a short diphthong, because [æ] is often treated as a "lax" vowel, and lax vowels are usually short in English. The question is whether this vowel is indeed short. Phonetically, [æ] is clearly a long vowel (Peterson and Lehiste 1960: 701). Phonologically, the New York [æ] undergoes "tensing" in such an environment (Benua 1995), and tense vowels in English are phonologically long, in part because they attract stress (Halle and Vergnaud 1987; Hayes 1995). Thus, there is good evidence that the New York [æᵊ] is not a short diphthong but a long one, which can be decomposed into two sounds.

Another reported example of short diphthongs is found in Gussmann's (2002) analysis of Icelandic. The argument is that the maximal rime size in Icelandic is VX (i.e. VV or VC); therefore, in a VVC rime, VV must be a short diphthong. It is worth noting that in VVC rimes, the C is typically a nasal. Such a case is not new. For example, Borowsky (1986) observes that in non-final English syllables the rime size is mostly limited to VX, although some exceptions are found, such as *pumpkin*, whose first rime is VCC, and *fountain*, whose first rime is VVC. In such rimes, the vowel is followed by a nasal, which can form a nasalized vowel (Duanmu 2008). Thus, the first rime in *pumpkin* is [ʌ̃p] and that in *fountain* is [aũ], an analysis independently proposed before (Malécot 1960; Bailey 1978; Fujimura 1979; Cohn 1993).

The examples show that different decisions on the granularity of segmentation can yield different consonants and vowels. When we examine databases of sound inventories, therefore, we should not take reported inventories at face value. Instead, we should be aware of the range of ambiguities and alternative interpretations. In addition, we should be cautious in drawing certain kinds of generalization, such as the number of consonants and vowels in a language, the average sizes of phoneme inventories across languages, or the total number of distinct sounds in all languages. Moreover, the size of a sound affects its feature analysis, too. For example, if [ʘx] is a single sound in !Xóõ, its feature analysis would be quite complicated. On the other hand, in a "cluster analysis" (Traill 1985: 208–11), [ʘx] is made of two sounds, a bilabial click [ʘ] and a velar fricative [x], and their feature analyses would be much simpler.

The granularity problem is related to what Chao (1934: 371) calls "under-analysis" and "over-analysis." In under-analysis, one treats what are "recognizably compound sounds" as a single sound, such as treating the affricate [ts] or [ts^h] as a single consonant. In over-analysis, one treats what is "one homogeneous sound" as two sounds, such as treating the American English vowel [ɚ] as [ə] plus [r]. But the terms "under-analysis" and "over-analysis" imply that (i) there is a "proper" analysis and (ii) we know what it is. In other words, Chao seems to assume that we already know what a sound is. The assumption is not obvious, as we shall see next.

1.1.3 *Defining sounds by time*

Given the assumption that speech can be segmented into a sequence of sounds, the simplest definition of sounds is as in (1), with examples in (2), where X is a time unit (McCarthy 1979; Pulleyblank 1983; Clements and Keyser 1983; Levin 1985). I use "features" as a cover term for phonetic properties (such as articulatory gestures), to be elaborated on later.

(1) Sounds defined by time
A sound is a set of compatible feature values in one time unit.

(2) Sample representations of sounds

<table>
<tr><td>Single sounds</td><td></td><td>Two-sound units</td><td></td><td></td></tr>
<tr><td>Simple</td><td>Complex</td><td>Diphthong</td><td>Long vowel</td><td></td></tr>
<tr><td>X</td><td>X</td><td>XX</td><td>XX</td><td>Timing tier</td></tr>
<tr><td>|</td><td>/\</td><td>| |</td><td>\/</td><td></td></tr>
<tr><td>k</td><td>kp</td><td>a i</td><td>i</td><td>Gestures</td></tr>
<tr><td>[k]</td><td>[k͡p]</td><td>[ai]</td><td>[iː]</td><td>Transcription</td></tr>
</table>

The notion of time is both phonetic and phonological. Phonetically, one sound is shorter than two sounds of comparable nature and in comparable context. For example, a short vowel (which takes one time unit) ought to be shorter than a long vowel or a diphthong (which take two time units each). Phonologically, a long vowel or a diphthong occupies two rime positions, similar to a short vowel plus a consonant. This can be demonstrated in some well-studied languages, such as English and Chinese. In addition, a long sound often shows different phonological behavior than a short one. For example, a long vowel tends to attract stress but a short one does not (Prince 1990; Hayes 1995). Similarly, a vowel may be lengthened when it carries a contour tone but not when it carries a level tone (Ward 1944).

Phonetic duration can be influenced by various factors though, in particular the phonotactics of a language. For example, an English vowel is shorter before a voiceless consonant than before a voiced one (House and Fairbanks 1953; Peterson and Lehiste 1960); thus, [ɪ] is shorter in *fit* than in *fig*. In addition, high vowels are shorter than non-high vowels. More striking cases have also been reported, in which the duration of a two-consonant cluster is similar to that of a single consonant. For example, Browman and Goldstein (1986: 233) report that in English, the temporal gesture of labial closure is the same for [p b m mp mb] (as in *capper*, *cabber*, *cammer*, *camper*, and *camber*), "regardless of whether the consonantal portion is described as a single consonant (/b/, /p/ or /m/) or as a consonant cluster (/mp/ or /mb/)." Similarly, Maddieson (1989) reports that, while prenasalized stops are found to be longer than a single consonant in Luganda (Herbert 1975; 1986), they have similar timing patterns to those of single consonants in Fijian. Such cases can sometimes be explained by the phonology of the given language though. For example, it can be argued that a prenasalized stop in Luganda is a true consonant cluster (Herbert 1975; 1986). In contrast, the Fijian consonant inventory has [p t k], [m n ŋ], and [^{m}b ^{n}d $^{\text{ŋ}}$g], but no [b d g] (Dixon 1988: 13); therefore, it is possible that the Fijian [^{m}b ^{n}d $^{\text{ŋ}}$g] are in fact single consonants [b d g] (Herbert 1975; 1986). In English, vowels are nasalized before a nasal; therefore, the nasal closure itself becomes redundant if the following consonant can indicate its place of articulation. In other words, the closure pattern observed by Browman and Goldstein (1986) can be described by a rule [VNC] → [ṼC], where [NC] is a homorganic cluster.

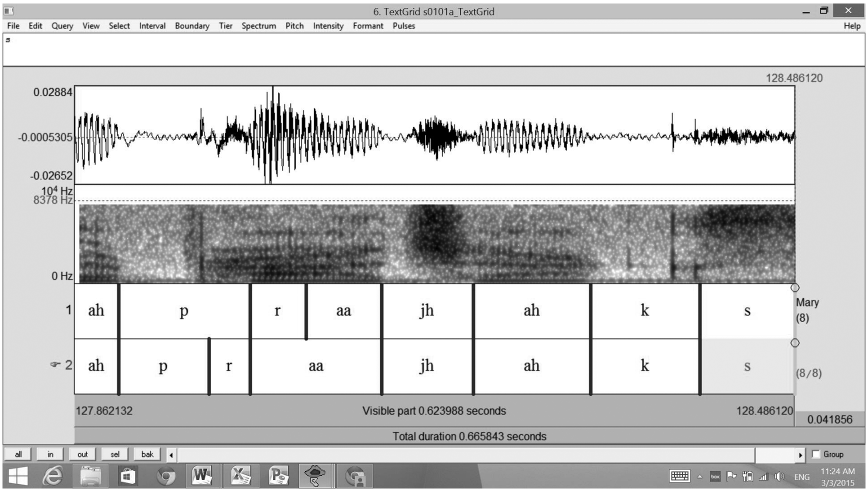

FIG. 1.1 Segmentation of '(the) projects' [prɑʤəks] by speaker s0101a of the Buckeye Corpus (Pitt et al. 2007), displayed in Praat (Boersma 2001). The upper tier shows the original segmentation of the Buckeye Corpus. The lower tier shows an alternative segmentation. Differences between the two segmentation decisions are discussed in the text.

Another source of variation in the duration of speech sounds may come from inconsistent decisions on segmentation. Consider the durational difference between [ɑ] and [ə] in Fig. 1.1, displayed in Praat (Boersma 2001), using the data of speaker s0101a in the Buckeye Corpus (Pitt et al. 2007).

In the original segmentation (upper tier), [k] does not include the aspiration portion, but [p] does. In addition, the location of [r] is questionable, since it is fully voiced, has greater intensity than the stressed vowel [ɑ], and does not match the auditory impression of the segment. Moreover, the stressed vowel [ɑ] is shorter than the unstressed vowel [ə]. In the alternative segmentation (lower tier), [p] and [k] have the same end point (right after release but not including aspiration). In addition, [r] is voiceless, in agreement with a well-known rule that "/w, r, j, l/ are at least partially voiceless when they occur after initial /p, t, k/" (Ladefoged and Johnson 2011: 73). The phonetics of [ɑ] and [ə] now seem more reasonable, too, where [ɑ] has greater intensity and is longer than [ə].

The idea of defining sounds by time has been proposed before. For example, Trubetzkoy (1969: 58) proposes that the duration of a sound should "not exceed the duration of the realization of other phonemes that occur in a given language." However, Trubetzkoy does not always stick with the time requirement. For example, he considers diphthongs, triphthongs, and long vowels in English to be single vowels (Trubetzkoy 1969: 118). In contrast, in the present analysis, a diphthong takes two time units and counts as two sounds, so does a long vowel.

A sound can be simple, such as [k], or complex, such as [k͡p] ([k] and [p] pronounced simultaneously). Now if [k] and [p] can form a complex sound [k͡p], can [a] and [i] form a short diphthong? The answer depends on the meaning of "compatible feature values" in (1), which can also distinguish possible and impossible complex sounds. There are two cases where feature values are compatible. First, the features involve independent articulators, such as [+stop] for Dorsal and [+stop] for Labial in [k͡p]. Second, the features are different, such as [+high] and [+back] for Dorsal. Two feature values are incompatible if they involve the same feature for the same articulator, such as [+nasal] and [–nasal] for Velum in [nd] and [–high] and [+high] for Dorsal in [ai]. In Chapter 7 we shall see that feature compatibility and complex sounds can be defined by the No Contour Principle, first proposed by Duanmu (1994).

1.1.4 Phonemes and allophones

A phoneme is a contrastive sound unit in a language (conventionally delimited by slashes), and allophones are different realizations of it (conventionally delimited by square brackets). For example, in English the phoneme /t/ can be realized as [t] in *stop* or [t^h] in *top*, and the phoneme /æ/ can be realized as [æ] in *pad* or [æ̃] in *pan*. In this sense, allophones, such as [t t^h], are concrete sounds, while a phoneme, such as /t/, is an abstract representation of them.

We are interested in possible contrasts in all languages. For example, we would like to know not just the fact that [t t^h] belong to the same phoneme in English, but whether [t] and [t^h] can ever distinguish words in other languages. Therefore, our focus is on concrete sounds and possible contrasts among them.

Now if we use data from phonemic inventories, two questions arise. First, if a phoneme is an abstract entity, how do we know its phonetic content? The answer is that, as a matter of practice, the symbol for a phoneme is usually the same as that of one of its allophones. For example, the English phoneme /t/ and its allophone [t] share the same symbol; so do the phoneme /æ/ and its allophone [æ]. Indeed, some linguists propose that a phoneme should always be represented by one of its allophones (Ruhlen 1976). Whether all linguists follow this practice or not, it is safe to say that most symbols in a phoneme inventory database correspond to one of its allophones and represent concrete sounds.

The second question is, when we see a phonemic symbol, how do we know what its allophones are? For example, when we see the phoneme [t] in English, how do we know it has the allophone [t^h]? The answer is that, if two allophones A and B never contrast in any language, they need not be distinguished. If A and B do contrast in some language, we will come across them when we examine that inventory. By examining all inventories available, we are able to identify all possible contrasts, even without allophonic information from any language.

1.2 Features and contrast

Features (also called "distinctive features") are used to serve three main purposes. First, they are used to represent a categorical difference between two sounds that can contrast in a language (i.e. distinguish words in that language). Second, they are used to account for sound classes (often called "natural classes"). Third, they are used to describe the phonetic properties of sounds.

1.2.1 Contrast

There is no question that every contrast must be represented. Therefore, I define features in (3), define contrast in (4) and give examples in (5) and (6).

(3) Features defined by contrast
A feature represents a minimal contrast between two sounds.

(4) Contrast
A contrast is the phonetic difference between one sound pair that can distinguish words in a language.

(5) Examples of contrast in English

Word pair	*Contrast?*	*Different sound pairs*
sip-zip	Yes	one pair [s]-[z]
sly-cry	No	two pairs [s]-[k] and [l]-[r]

(6) Degree of contrast

Minimal	[s]-[z]	*sip-zip*
Minimal	[s]-[f]	*see-fee*
Non-minimal	[s]-[v]	*set-vet*

In (5), the difference between *sip-zip* is provided entirely by that between [s]-[z], which confirms that [s] and [z] can offer a contrast. On the other hand, the difference between *sly-cry* is provided by [s]-[k] and [l]-[r]; therefore, the word pair does not show whether [s]-[k] or [l]-[r] can provide a contrast alone.

The present notion of features shares some similarity with what Trubetzkoy (1969) calls "oppositions." However, in the present definition, a feature refers to a difference between two sounds, whereas for Trubetzkoy an opposition can also refer to a difference between other phonological units, such as between the lengths of two syllables.

Since a contrast is found between a pair of concrete sounds, the difference between the pair can be defined phonetically, as can features. Such definitions allow us to determine the degree of differences or similarities between contrastive sound pairs. For example, the difference between [s]-[z] is the same as that between [f]-[v], commonly known as [voice]. Also, it can be shown that [voice] is minimal, in the sense that it cannot be further divided into two (or more) components each of which can itself be contrastive. It can be shown, too, that the difference between [s] and [f] is also minimal, commonly known as the "place" of articulation. On the other hand, it can be shown that the difference between [s] and [v] is not minimal, in the sense that it can be divided into two components, [voice] and "place," each of which can be contrastive by itself.

A contrast-based definition of features also means that non-contrastive phonetic differences are not represented by features. For example, the vowels of a child differ from those of an adult, but such differences do not distinguish words and so are not represented by features. Similarly, vowels of an average male differ from those of an average female, and such differences are not represented by features either.

The notion of contrast can be used not only to determine phonemes, as in (7), but also to define allophones, as in (8).

(7) Phonemes
If two sounds contrast (can distinguish words) in a language, they belong to different phonemes in that language.

(8) Allophones
Two sounds X and Y (which share some phonetic similarity) are allophones of the same phoneme in a language if and only if
a. X and Y do not contrast in this language, and
b. X and Y can distinguish words in another language.

Without (8b), a phoneme could have an infinite number of allophones. For example, each individual's [i] is different from that of another speaker, and every utterance of [i] by the same speaker is somewhat different. With (8b), the number of allophones of a phoneme is radically reduced. For example, a child's [i] and an adult's [i] are not treated as allophones.

But how do we know whether a difference between two sounds is potentially distinctive in another language? For example, can a normal vowel and a whispered one ever contrast in any language? Similarly, can [m] and [ɱ] ever contrast in any language? Such questions require an examination of large databases of sound inventories, which is the focus of the present study.

1.2.2 *Sound classes*

Features can also be defined by sound classes, which is shown in (9); a sound class is defined in (10).

(9) Features defined by sound classes
A feature represents a minimal similarity shared by members of a sound class.

(10) Sound class
A sound class is a set of sounds that behave alike in a phonological pattern.

Sound classes are often called "natural classes." However, as Mielke (2008) points out, many sound classes are in fact unnatural, in that there is no discernible common feature among the members. Therefore, he offers the term "phonologically active classes" instead. For brevity, I shall use the term "sound classes," without implying whether or not they are natural, or phonologically active.

The similarity among members of a sound class is sometimes easy to see. For example, [b d g] share the similarity of "voiced," "oral," and "stop," [b d g p t k] share the similarity of "oral" and "stop," and [b d g p t k m n ŋ] share the similarity of "stop" only. The similarity in the last set is minimal and involves one feature. The similarities in the first two sets are not minimal and involve more than one feature. Sound classes have been used not only to establish features but also to establish hierarchical relations among features, where all features are organized into a tree structure (called "feature geometry"). For example, Clements (1985) groups [nasal], [continuant], and [strident] under a higher node called "manner," and Sagey (1986) and Halle (1995) group "labial," "coronal," and "dorsal" under a higher node called "place."

However, it is not always clear what constitutes a sound class, or what feature(s) the members of a sound class share. Let us consider some examples from P-base (Mielke 2004–7), a comprehensive collection of sound classes. For English, [m n ŋ] are listed as a sound class, because they can nasalize a preceding vowel. There are two questions, though. First, is there ever a contrast between a nasalized vowel and a non-nasalized one before a nasal consonant? If not, should vowel nasalization, in this context, be treated in terms of features? Second, suppose [m n ŋ] do form a sound class. How many features do they share? It may appear that they share four, [+nasal, +stop, +voice, +sonorant]. However, it is possible they share just one feature [+nasal], from which we can predict the other three, the latter being redundant (see Chapter 6 on feature specificstion).

Next we consider the plural suffix in English, which can be realized as [s] (as in *cats*), [ɪz] (as in *buses*), and [z] (as in *fans*). P-base considers there to be three sound classes, one preceding [s], one preceding [ɪz], and one preceding [z]. However, it can be argued that [z] is the default form and only two classes are needed: one triggering [ɪz] and one triggering [s]. In addition, as seen above, even if we have determined a sound class, it is not easy to tell what defines it. For example, what is the feature that defines the class that precedes [s]? Is it [–voice] (McMahon 2002) or is it [+aspirated] (Iverson and Salmons 1995; Beckman et al. 2009)?

As a final example, consider the set of English consonants that can occur before [l] in word-initial position, which are [p b k g f v s ʃ]. P-base lists this set as a sound class, but questions can be raised. For example, why are [t] and [d] missing from it? The reason seems to be dissimilation (both [l] and [t d] are coronals), but why are [s] and [ʃ] allowed (since they also involve a coronal component)? If dissimilation does not hold for fricatives, why are [z] and [ʒ] missing? An alternative answer is that word-initial [s] is a special case, as noted in the literature, as is word-initial [ʃ] (Duanmu 2008). If we exclude [s] and [ʃ], the remaining set is [p b k g f v], which is a lot easier to define.

The examples show that analyses of sound classes can be complicated, even for well-studied languages. To see the extent of uncertainty, consider the forty sound classes listed by P-base for English. Of those, twenty-seven are based on distributions at word edges, which are not always a reliable test for sound classes. Of the remaining thirteen, seven are controversial in various ways. Thus, Mielke (2008) concludes, correctly in my view, that many apparent sound classes are unnatural, as are features construed from them. A similar point has been made by Steriade (1999) and Hayes (1999), who argue that some phonological patterns are not based on feature-defined sound classes but can be better accounted for either by function-oriented phonetics or by constraint interaction. This is not to say that sound classes are not worth studying. Rather, of the many sound classes collected in Mielke (2008), many have not been carefully examined before, if at all. Clearly, they call for a separate study, which is beyond the scope of the present one.

1.2.3 Phonetic properties

Features are also used to represent the phonetic properties of sounds. For example, [f] can be represented as a labiodental voiceless fricative, where "labiodental" refers to an articulatory gesture (moving the lower lip against the upper teeth), "voiceless" refers to an articulatory gesture (lack of vocal-fold vibration) or its acoustic effect (lack of periodic sound wave), and "fricative" refers to an articulatory gesture (a type of closure) or its acoustic effect (presence of aperiodic sound wave). This approach has a long tradition, codified by the IPA (International Phonetic Alphabet) chart, where sounds are arranged in columns and rows according to their articulatory features.

In the phonetic definition, a feature is a phonetic dimension, which indicates an articulatory gesture or its acoustic or perceptual effect (Jakobson et al. 1952; Abercrombie 1967; Chomsky and Halle 1968; Ladefoged 1972). Each dimension in turn is divided into two (or more) values or degrees. Some examples are shown in (11).

(11) Features (phonetic dimensions) and their values

Feature	*Value*
Voice	Voiced, voiceless
Aspiration	Aspirated, unaspirated
Backness	Front, (central), back
etc.	

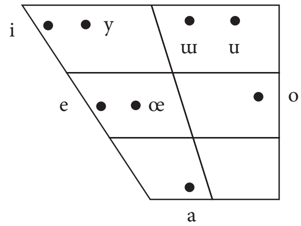

FIG. 1.2 Phonetic properties of Turkish vowels (Zimmer and Orgun 1992: 44). Reproduced with permission from Cambridge University Press.

The features "voice" and "aspiration" have two values each. The feature "backness" (of the tongue) has been divided into two values (front–back) or three (front–central–back).

However, it is not always clear which phonetic properties are relevant, or how many degrees a feature should be divided into. Consider Turkish vowels, whose phonetic and phonological properties are shown in Fig. 1.2 and in (12).

(12) Phonological features of Turkish vowels

	Front		Back	
	–round	+round	–round	+round
+high	i	y	ɯ	u
–high	e	œ	a	o

Phonetically, the Turkish [a] is low and central (or front), but phonologically, it is a back vowel, and in terms of height it belongs to the same class as [e œ o] (Lewis 1967). Without phonological data, one might interpret Fig. 1.2 literally and arrive at the analysis in (13).

(13) A phonetic interpretation of Turkish vowel features based on Fig. 1.2

	Front		Back	
	–round	+round	–round	+round
high	i	y	ɯ	u
mid	e	œ		o
low	a			

However, there is clear evidence that (12) is a better analysis. Consider the alternation of the plural suffix, shown in (14).

(14) Alternation of the plural suffix in Turkish

Preceding vowel	Plural suffix
Front: [i y e œ]	[ler]
Back: [ɯ u a o]	[lar]

When the preceding vowel is one of [i y e œ], the plural suffix is [ler], and when the preceding vowel is one of [ɯ u a o], the plural suffix is [lar]. The pattern is accounted for if [i y e œ] are front vowels and [ɯ u a o] are back vowels, so that the suffix assumes the same backness as the preceding vowel. In addition, the pattern suggests that [e a] only differ in backness and not in height. If [a] is a front vowel, as shown in (13), there is no reason why it changes its height to [e], based on the backness, not height, of the preceding vowel. The data also show that the backness difference between [i y] (or between [ɯ u] or [e œ]) has no effect. Clearly, phonetic properties require reinterpretation when we consider patterns from contrast and sound classes.

1.2.4 *Summary*

Ideally, features obtained from different criteria (contrast, sound classes, and phonetic properties) coincide with one another (Halle 1962; 1995); but the ideal remains to be confirmed.

In this study I shall focus on contrast-based features, with some reference to their phonetic properties but only occasional references to sound classes. There are several reasons. First, a contrast-based feature system is the least controversial, because all linguists agree that every contrast must be distinguished. Second, our data mostly consist of phonetic symbols, which have encoded a substantial amount of phonetic information. In particular, the transcription of a sound indicates its key phonetic properties in relation to those of other sounds. For example, [i] is always the highest vowel in a language and [e] is always lower than [i]. Third, a contrast-based feature system must be recognized independently, no matter what patterns emerge from sound classes. Specifically, if two sounds contrast in a language, they must be distinguished, with or without information from sound classes. Finally, a contrast-based feature system can help us interpret phonetic properties. For example, if no language makes a three-way contrast in backness, we must reinterpret what appears to be a central vowel as either front or back.

1.3 Cross-language comparison

Any study that compares sounds across languages is faced with a fundamental methodological question: Let X be a sound from one language and Y a sound from another language. How do we decide whether X and Y should be treated as the same sound or different sounds? Consider Fig. 1.3, which shows four vowels in German (dotted lines, Jørgensen 1969) and four in Norwegian (solid lines, Gamnes 1965), plotted by Disner (1983: 67).

While both languages use the symbols [i y e ø], the German vowels are systematically higher, and the German [y] is both higher and less front. Such small but systematic differences between languages are quite common (Ladefoged 1972; Disner 1983). What is the reason for saying that the Norwegian vowels are the same as those in German, beyond the fact that they are represented by the same phonetic symbols, probably for convenience? Should such differences be distinguished at all?

Consider another example, illustrated with the backness of the tongue, shown in Figure 1.4, where A–D are four vowels in two hypothetical languages, L1 and L2.

If we considered L1 alone, we may call A front and B back. Similarly, if we consider L2 alone, we may call C front and D back. Moreover, if all languages have only two degrees of backness, we can identify A with C, both being front, and

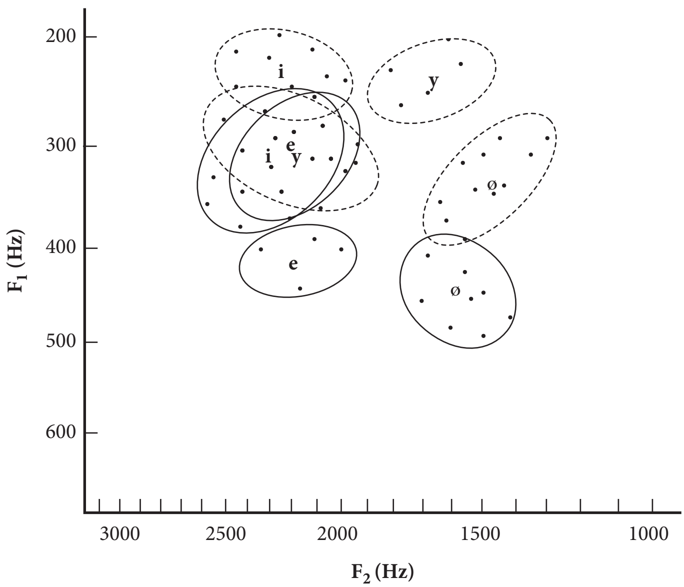

Fig. 1.3 Four vowels in Norwegian (solid line, 10 speakers) and German (dotted line, 6 speakers), from Disner (1983: 67).

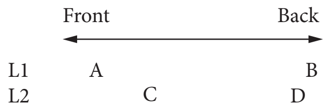

FIG. 1.4 Backness of four vowels, A–D, in two hypothetical languages, L1 and L2. If all languages have only two degrees of backness, we can identify A with C, both being front, and B with D, both being back.

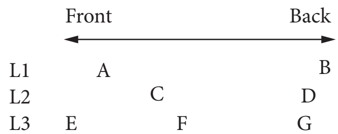

FIG. 1.5 Backness of seven vowels, A–G, in three hypothetical languages, L1–L3, where A–D are the same as those in Fig. 1.4. Given E and F, the analysis of C needs to be reconsidered: C seems to be closer to the central vowel F, rather than to the front vowel A or E.

B with D, both being back, even though there is some difference between A and C, or between B and D.

However, if there is a language L3 that has three degrees of backness, as shown in Fig. 1.5, we have to reconsider the analysis of C.

The vowels E, F, and G seem to be front, central, and back respectively. Given E and F, it seems more reasonable to consider C to be a central vowel, rather than a front vowel. Now what if a language has four degrees of contrast in backness? Do we need to change the analysis of C again? Clearly, cross-language comparisons can be quite uncertain, because patterns from other languages can often change our conclusions. Such questions have been noted before, e.g. by Lass (1984), who did not come up with a workable solution.

Linguists differ on how to address such questions. Let us consider three views. According to the first (e.g. Boas 1896; Joos 1957; Ladefoged 1972; 1992; Disner 1983; Port and Leary 2005), each language is different and should be analyzed on its own. It makes little sense to identify sounds in one language with those in another. Nor is it meaningful to ask how many sounds there are in the world's languages (unless we consider each language to have a different set of sounds, which makes the question trivial). For example, in L1 of Fig. 1.4, we can call A front and B back, and in L2 we can call C front and D back, but it makes little sense to identify A with C, because "front" in L1 does not mean the same as "front" in L2. However, this approach overlooks some important questions, such as the maximal number of contrasts in each phonetic dimension. In addition, this approach fails to appreciate the possibility that a mapping relation can hold between sounds that are not phonetically identical, such as A and C in Fig. 1.4, if no language has more than two degrees of contrast in backness.

According to the second view, sounds in different languages can be equated to each other, if we have a universal feature system (e.g. Trubetzkoy 1969; Jakobson et al. 1952; Chomsky and Halle 1968). However, it remains to be shown how such a feature system is discerned from inventory databases. In addition, it remains to be explained why there are small but systematic differences between languages.

According to the third view, features can be derived from physical landmarks in the vocal tract (the "quantal theory" of Stevens 1972; 1989). Therefore, at least some features can be identified (or equated to each other) across languages. It is unclear, though, how many degrees of contrast such a theory would predict for each phonetic dimension, and how well the predictions would fare against available data.

Despite the differences among the views, there is an empirical question that should be of interest to all: What is the maximal number of contrasts in each phonetic dimension? By finding the answer, we can evaluate current theories. For example, if the maximal number of contrasts in backness is two, then it is reasonable to say that backness can be compared across languages. If the maximal number is more than two, then the binary-feature theory becomes less attractive and cross-language comparison becomes less obvious. In this study, therefore, we shall examine the maximal number of contrasts in every phonetic dimension.

1.4 Adequacy of available data

This study uses two phoneme inventory databases. One is the UCLA Phonological Segment Inventory Database (UPSID, Maddieson and Precoda 1990), which contains 451 inventories. The other is P-base (Mielke 2004–7), which contains 628 inventories. Compared with the total number of languages in the world today, estimated to be 6,000 (Moseley 2010), the databases seem small. Therefore, one might ask whether they are adequate.

UPSID was compiled by selecting one language from each typological group. Therefore, it is a reasonable representation of the world's languages. P-base was compiled by collecting all inventories on which there is a published grammar book at the libraries of two large universities, Ohio State University and Michigan State University. It is, therefore, also a good representation of available data. Both databases were mainly based on sources in English.

Some linguists are optimistic with regard to how much we already know. For example, Ladefoged and Maddieson (1996: 1–2) offer the following upbeat statement:

> We believe that enough is now known to attempt a description of the sounds in all the languages of the world... 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.

Besides the issue of coverage, other questions have been raised (e.g. Lass 1984; Simpson 1999; Vaux 2009). In particular, there is the problem of granularity of segmentation in phonemic analysis. In addition, there is an issue of typographic convenience, as noted by Ruhlen (1976). For example, the IPA symbol [a] is intended to be a low front vowel and [ɑ] a low back vowel, but when a language does not have both, [a] is often chosen instead of [ɑ]. Similarly, for those who use a typewriter (or a simple keyboard), [ɨ] may be favored over [ɯ], because [ɨ] can be made with a key combination ([i] with strikethrough) but [ɯ] cannot. Moreover, different analyses may choose different symbols to represent phonemes. For example, in P-base, Spanish has [b d ɡ], but in UPSID they are represented as [β ð ɣ]. Such problems may be an issue if we are interested in the frequencies or markedness of sounds (Basbøll 1981; Brakel 1983; Calabrese 1995; Rice 2006; Clements 2009), but not if we are interested in contrasts among different sounds. For example, whether the Spanish sounds are [b d ɡ] or [β ð ɣ], we may need to distinguish all of them, if some languages have all of them. Similarly, whether a low vowel is [a] or [ɑ] in a given language, as long as some language has both, we can capture the contrast. Moreover, as will be discussed in Chapter 2, our method aims to identify all inventories that appear to contain unusual contrasts, and every such case will be examined manually.

1.5 Summary

I have reviewed controversies in defining sounds and features, difficulties in using databases of phoneme inventories, and problems in cross-language comparisons. The discussion points to the need to determine the maximal number of contrasts in each feature (as a phonetic dimension). The result should be of interest to all parties in the theoretical debate.

2

Method

This chapter covers the method of the present study, in particular the databases used and the principles that guide the procedure. I also illustrate the approach with examples.

2.1 Data

The data consist of two sound inventory databases, UPSID (the UCLA Phonological Segment Inventory Database, Maddieson 1984; Maddieson and Precoda 1990) and P-base (Mielke 2004–7). They are both publically available in electronic form.

The selection method of UPSID was to choose one language per typological group. The selection method of P-base was to include all "language grammars (written in English) available in the Ohio State University and Michigan State University library systems." Every selected grammar "is based on data collected while the language was still living" (Mielke 2008: 48). A comparison of UPSID and P-base is shown in Table 2.1, where a token is an occurrence of a phoneme (or its phonetic transcription) in an inventory.

Most sources of P-base are books, because its aim is to collect not only phoneme inventories but also sound classes. In contrast, many sources of UPSID are short articles that offer little beyond a phoneme inventory. In addition, while both databases converted original transcriptions into a consistent system, UPSID was quite

TABLE 2.1 A comparison of UPSID and P-base

	UPSID	P-base
Inventories	451	628
Phoneme tokens	13,966	19,959
Coverage	One per typological group	All grammar books (written in English)
Type of sources	Articles and books	Mostly books
Reinterpretation	Liberal	Minimal

A Theory of Phonological Features. First edition. San Duanmu.

liberal at reinterpreting the original transcriptions, whereas P-base aimed at preserving the transcriptions of the original authors.

Based on complete identity of language names, there is an overlap of about 100 inventories between the two databases. Therefore, the total number of different inventories in the two databases is close to 1,000 or 17 per cent of the world's languages, the latter being estimated to be 6,000 (Moseley 2010).

2.2 Guiding principles

As mentioned in Chapter 1, the goal of this study is to determine a minimally sufficient feature system that can distinguish all consonants and vowels in the world's languages, based on the data in UPSID and P-base. I shall follow three principles, which I call the Principle of Contrast, Maxima First, and Known Feature First, defined in (1)–(3) and illustrated below.

(1) The Principle of Contrast
   a. If two sounds A and B can contrast in any language, they must be distinguished by at least one feature.
   b. If two sounds A and B never contrast in any language, they need not be distinguished by a feature.

(2) Maxima First
   a. First, search through all languages in order to determine the maximal number of contrasts in each phonetic dimension. When all dimensions have been examined, we obtain a maximal feature system.
   b. Then, interpret each sound of a language in terms of the maximal feature system.

(3) Known Feature First
   Unless evidence requires otherwise, use known features first before introducing a new feature (or a new feature combination) to represent a contrast.

The Principle of Contrast in (1) is a basic assumption in phonology and is noncontroversial. It can be seen, too, that (2) and (3) follow from (1), to be explained below.

2.2.1 *Principle of Contrast*

Since contrast is based on words of a given language, the Principle of Contrast is commonly used in the analysis of individual languages. For example, according to International Phonetic Association (1999), in the "broad transcription" of a language, non-contrastive differences (known as allophonic variations) are ignored.

However, if the Principle of Contrast is limited to individual languages, we face a problem when we do "narrow transcription" of a language (which captures both contrastive and non-contrastive differences), or when we compare sounds from different languages. For example, in the narrow transcription of English, we typically distinguish unaspirated [t] as in *stop* from aspirated [t^h] as in *top*. But what if someone has a slightly longer aspiration in the [t^h] of *top*? Should we transcribe it as [t^{hh}]? What is the reason for doing so, or not doing so? What if someone has slightly less aspiration in the [t^h] of *top*? Clearly, unless we have a principled answer, narrow transcription is both impossible and arbitrary: There are infinitely many non-contrastive details that are impossible to capture, and any decision on how many details to capture seems quite arbitrary.

A similar problem arises when we compare languages or dialects. For example, Ladefoged (1992) observes that, in American English (by speakers of California), the tongue tip is visible in [θ] (as in *thin*), whereas in British English (by speakers of Southern England), the tongue tip is not visible. Similarly, Disner (1983) observes that the [i] of German speakers has a higher tongue position than [i] of Norwegian speakers. Should such differences be captured in transcription? Clearly, between languages and dialects, there are infinitely many small differences that are observable; but unless we have a theory of what should or should not be represented, phonetic transcription is both impossible and arbitrary.

Our definition of the Principle of Contrast extends its use from individual languages to cross-language comparisons. The extension should not be controversial, since it makes the principle easier to falsify and the theory stronger. Our definition also offers a solution to the problem of transcription. In the narrow transcription of a language, we should only represent differences that are contrastive in another language, not differences that are not contrastive in any language. For example, [t] and [t^h] are contrastive in Hindi and should be represented in the narrow transcription of English. On the other hand, there is no known contrast between [t^h] (regular amount of aspiration) and [t^{hh}] (extra amount of aspiration); therefore, there is no need to distinguish the two in any language.

The problem in cross language comparison has a similar solution. For example, if the difference between protruded [θ] (in American English) and non-protruded [θ] (in British English) is contrastive in any language, the two sounds must be distinguished. Otherwise, there is no need to represent the difference. Similarly, consider "linguo-labial" (or "apical-labial") consonants, which have been reported in some languages (Postal 1968; Tryon 1976; Maddieson 1987; Ladefoged and Maddieson 1996; Olson et al. 2009). A linguo-labial consonant is similar to an interdental consonant, except that in the former the tongue tip is further forward so that it touches the upper lip. Acoustically, linguo-dentals are similar to interdentals, too (Maddieson 1987: 26). Should we distinguish linguo-dentals from interdentals or consider them to be variants of the same sounds? The answer again lies in contrast: If linguo-dentals and interdentals can contrast in any language, they must be

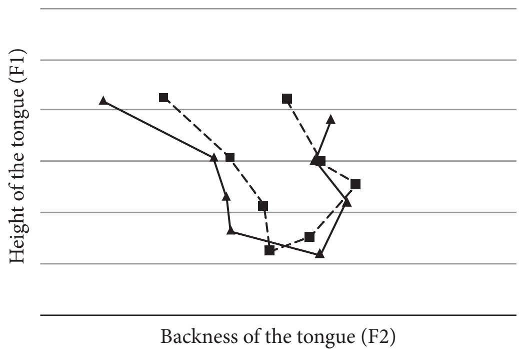

FIG. 2.1 [i ɪ ɛ æ ɑ ɒ ʊ u] by two female speakers of American English, measured by San Duanmu. Neither speaker thought the other had a different accent.

distinguished. If linguo-dentals and interdentals never contrast in any language (Ladefoged and Maddieson 1996: 40), they need not be distinguished.

This is not to say that non-contrastive differences are completely ignored. Rather, as I shall suggest in Chapter 6, there are explanations for non-contrastive differences, without resorting to feature differences.

There is good evidence for the Principle of Contrast. Consider the examples in Figs 2.1–2.3, where the height and backness of the tongue are reflected by F1 and F2 values.

Figure 2.1 shows eight vowels by two female speakers of American English. It can be seen that some corresponding vowel pairs are quite different phonetically, yet their differences are ignored, since neither speaker considered the other to have any accent. Similarly, Figure 2.2 shows that, while vowels of speakers of the same language differ a lot, the differences are ignored and the ten vowels remain distinct. Figure 2.3 shows that even for the same speaker, what are heard as [i u ɑ] by phonetic transcribers in fact vary a lot, which shows again that non-contrastive phonetic differences can be (and are) ignored.

2.2.2 *Maxima First*

The idea of Maxima First is not entirely new. For example, one might consider the IPA chart to be a maximal system of contrast, which offers a sufficient number of contrasts in each phonetic dimension. However, the IPA is the result of piecemeal development, not of a coherent theory of contrast. In particular, it is a system for representing conceivable differences, rather than a system for representing contrasts. As a result, the IPA offers far too many distinctions than warranted by contrast. For example, consider backness among unrounded high vowels. The IPA distinguishes nine degrees of backness, [i̟]-[i]-[i̠]-[ɨ̟]-[ɨ]-[ɨ̠]-[ɯ̟]-[ɯ]-[ɯ̠]. It can be shown that no

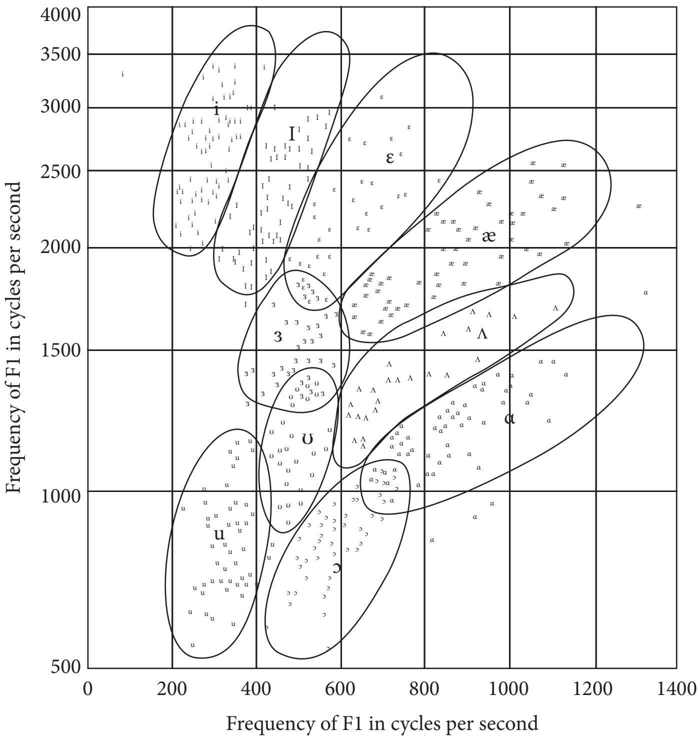

Fig. 2.2 Ten vowels by seventy-six American English speakers. The X-axis shows F1 values and the Y-axis F2 values. Reprinted from Peterson and Barney (1952: 182). Copyright 1952, Acoustic Society of America.

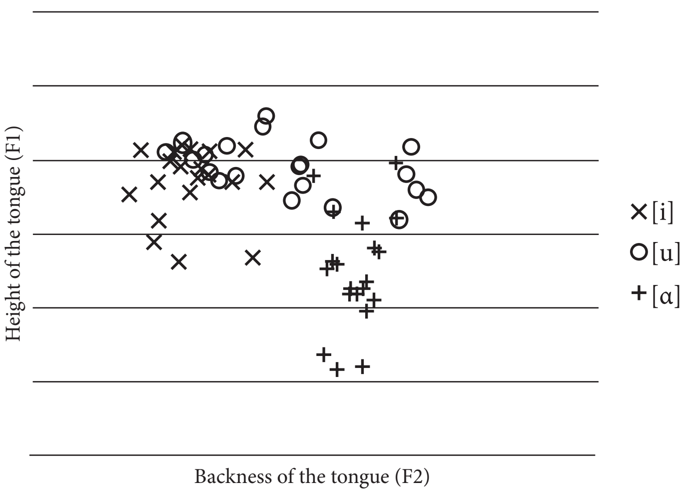

Fig. 2.3 Contextual variation of three vowels (twenty tokens each) by one female speaker of American English. The vowels were narrowly transcribed as [i], [ɑ], or [u], spoken by speaker s0101a from Columbus, Ohio, in the Buckeye Corpus (Pitt et al. 2007; measured by San Duanmu).

language has so many degrees of contrast in backness. Therefore, to establish a minimally sufficient system of contrasts, we need to determine the maximal number of occurring contrasts in each phonetic dimension, by searching through all available inventory databases.

Maxima First interacts with Known Feature First, in that when a language seems to show a larger than expected number of contrasts in a feature, we need to ask whether the contrasts can be represented with two (or more) known features.

2.2.3 Known Feature First

The main function of Known Feature First is to minimize redundancy in representation: If a difference is already represented, we do not need to represent it again in another way. For example, consider the difference between [ə] and [ʌ] in English, shown in (4).

(4) Representing the difference between [ə] and [ʌ] in English

	Feature	Stress
[ə]	Central	Unstressed
[ʌ]	Back	Stressed

In transcriptions that distinguish [ə] and [ʌ], [ə] appears to be central and [ʌ] back. However, [ə] is an unstressed vowel and [ʌ] a stressed one. Since the distinction is already represented by stress, there is no need to represent it again by a feature (or an extra degree in the feature backness).

Next we consider vowel length. There are two ways to represent the difference between a long vowel and a short one, shown in (5).

(5) Representing vowel length

Property of sound		*Property of structure*	
Short	*Long*	*Short*	*Long*
[i]	[i:]	[i]	[i]
		|	/\
		X	XX

In some analyses, such as Kashmiri in UPSID, vowel length is seen as the property of the sound, which needs to be represented by a feature. In other analyses, vowel length is represented by structure, i.e. timing slots (or mora count), where a short vowel has one and a long vowel has two. If timing slots are independently motivated (Goldsmith 1976; McCarthy 1979; Pulleyblank 1983; Levin 1985), we can represent vowel length in all languages with timing slots, with no need for an additional length feature.

Finally, consider vowel height (or factors that affect vowel height). Two options are shown in (6).

(6) Representing vowel height (ATR = advanced tongue root)

	1 feature	*2 features*
[i]	high 1	+high, +ATR
[ɪ]	high 2	+high, −ATR
[e]	high 3	−high, +ATR
[ɛ]	high 4	−high, −ATR

If we use one feature, we need a four-way contrast in height. If we use two features, [high] and [ATR] (Jakobson and Halle 1964; Ladefoged 1964; Stewart 1967; Halle and Stevens 1969; Perkell 1969; 1971; Lindau 1979; Wood 1982; Vaux 1996; Kenstowicz 2009; Allen et al. 2013), we need a two-way contrast in each. It can be shown that in some languages, such as Kinande (Kenstowicz 2009) and Yoruba (Allen et al. 2013), two features are required. Therefore, unless there is evidence otherwise, we can use the same representation for other languages, without using a four-way contrast in height.

Besides minimizing new features, Known Feature First can also minimize new feature combinations. For example, consider the contrast between [l] and [ɬ]. There are two ways to represent the difference, shown in (7), where [lat], [asp], and [fric] are abbreviations for [lateral], [aspirated], and [fricative] respectively.

(7) Representing the contrast between [l] and [ɬ]

	[l]	[ɬ]
Analysis 1	[+lat, +voice, −asp]	[+lat, −voice, +asp]
Analysis 2	[+lat, +voice, −asp, −fric]	[+lat, −voice, +asp, +fric]

Laterals are normally sonorant, and sonorant sounds can be voiceless and aspirated (or breathy), such as voiceless nasals or breathy vowels. Therefore, the feature combination in Analysis 1 is not controversial. In contrast, in Analysis 2, not only is [fricative] redundant, but the combination of [+lateral] and [+fricative] is new. Therefore, unless there is evidence otherwise, we should prefer Analysis 1.

2.2.4 *Summary*

I have introduced three principles, of which the Principle of Contrast is the most important, from which the other two follow. For example, suppose there are two languages L1 and L2, where L1 has [i] but not [i̠], and L2 has [i̠] but not [i]. How do we decide whether the two sounds are the same or different? Based on the Principle of Contrast, we would search through all inventories and see if [i] and [i̠] contrast in any inventory.

Now if we find a language that has both [i] and [i̠], then we are sure they are different. But is [i̠], whose tongue position is further back than [i], different from [ɨ], whose tongue position is further back still? On the basis of the Principle of Contrast, we need to find out whether there is a language in which [i̠] contrasts with [i] to one side and [ɨ] to the other. In other words, we need to find out whether any language has a three-way contrast in [i]-[i̲]-[ɨ]. Clearly, soon we would be asking for the maximal number of contrast in backness, and this is what Maxima First is for. In other words, Maxima First follows from the Principle of Contrast.

Known Feature First follows from the Principle of Contrast, too. As mentioned above, the main function of Known Feature First is to minimize redundancy in representation. Consider [ə] and [ʌ] in English again, repeated in (8).

(8) Representing the difference between [ə] and [ʌ] in English

	Feature	Stress
[ə]	Central	Unstressed
[ʌ]	Back	Stressed

Since the difference between [ə] and [ʌ] can be represented by stress, and since stress is independently motivated, there is no additional contrast between [ə] and [ʌ]. And since features are based on contrast, when there is no contrast, there is no basis for using a backness feature here, or introducing a new degree in backness.

2.3 Interpreting transcription errors

It is hardly possible to build a large inventory database that is error-free. Some errors come from original sources that are hard to verify. Some errors come from non-optimal analyses. Some errors come from clerical oversight. Such errors have been pointed out before (Simpson 1999; Vaux 2009) and are thought to undermine the value of inventory databases. Fortunately, given the Principle of Contrast, many errors can be spotted and properly interpreted, leaving little adverse effect. As an example, consider the transcription of [d] and [ð]. A number of possible cases are shown in (9).

(9) Transcriptions and contrast between [d] and [ð]

Languages	Transcription	Comment
L1	[d]	Has [d] only; no error
L2	[ð]	Has [ð] only; no error
L3	[d] → [ð]	Has [d] only, miswritten as [ð]
L4	[ð] → [d]	Has [ð] only, miswritten as [d]
L5	[d (ð)]	[d ð] are allophones; [d] is chosen
L6	[(d) ð]	[d ð] are allophones; [ð] is chosen
L7	[d], [ð]	[d] and [ð] contrast; no error

L1 and L2 have no error, but they cannot tell us whether [d] and [ð] are the same or different, since neither language has a contrast between the two sounds. However, L7 offers the answer, where [d] and [ð] are different, because they contrast there. In L3 [d] is miswritten as [ð], and in L4 [ð] is miswritten as [d]; but the errors are limited to the given languages, and they have little impact on the general question of whether [d] and [ð] are the same or different, since we have the answer in L7 already. In L5 and L6, [d ð] are allophones, and only one is chosen to represent the phoneme. Simpson (1999) argues, correctly, that it is not easy to decide which allophone ought to be chosen. In addition, he argues that choosing a single symbol to represent a set of allophones means a massive loss of phonetic details. However, as far as the representation of contrast is concerned, the problem has little effect. As long as there is a language like L7, we know that [d] and [ð] are different. And if there is no language like L7, [d] and [ð] need not be distinguished in any language. The same argument applies to any other pair of symbols.

2.4 Searching for maxima

To determine the maximal number of contrasts in each phonetic dimension, we follow the procedure in (10).

(10) Search procedure for the maximal number of contrasts
- a. Extract a complete list of distinct transcriptions (sound types).
- b. Group the list of sound types according to their phonetic properties.
- c. Look for the maximal number of contrasts in each phonetic dimension.
- d. Extract all cases that seem exceptional or controversial (such as a three-way contrast in any phonetic dimension, or features with very low frequencies).

For illustration, let us consider backness in UPSID. There are 269 vowel types (distinct vowel transcriptions). Following Maddieson (2005), let us make a distinction between basic vowels (what he calls "basic monophthongs") and other vowels. The former involve tongue positions and lip rounding only. The latter also involve length, nasalization, laryngeal features, diphthongs, and some other cases. In UPSID, there are thirty-eight basic vowels, shown in Table 2.2.

A three-way contrast in backness is shown by twelve triplets, which are listed in (11). In each triplet, the three sounds differ in backness only.

(11) Twelve contrastive triplets in backness, from Table 2.2
[i ɨ ɯ], [ɪ ᵻ Ɯ], [e̞ ə̞ ɤ̞], [e ə ɤ], [ɛ ɜ ʌ], [æ ɐ ɐ̠], [a a̠ ɑ]
[y ʉ u], [ʏ ɵ ʊ], [ø̞ ɵ̞ o̞], [ø ɵ o], [œ ɞ ɔ]

Next, we search through all inventories to identify languages that contain one or more of the triplets. Three languages are found, with one triplet each, shown in (12).

Table 2.2 Thirty-eight basic vowels found in UPSID; four empty positions represent vowels not found

	Front		Central		Back	
High	i	y	ɨ	ʉ	ɯ	u
High (lower)	ɪ	ʏ	ɪ̵	ʊ̵	Ɯ	ʊ
Mid (higher)	e̝	ø̝	ə̝	ɵ̝	ɤ̝	o̝
Mid	e	ø	ə	ɵ	ɤ	o
Mid (lower)	ɛ	œ	ɜ	ɞ	ʌ	ɔ
Low (raised)	æ		ɐ		ɐ̠	ɐ̠ʷ
Low	a		a̠		ɑ	ɒ

(12) *Language* *Triplet found*

Moro	[e ə ɤ]
Nimboran	[i ɨ ɯ]
Woisika	[a a̠ ɑ]

The three languages, and their sources, are then examined in order to evaluate the validity of the data.

2.5 Interpreting exceptions

To be exhaustive, we cast a wide net and define exceptions broadly. Any feature, or any extra degree of contrast in a feature, that is controversial is treated as an exception and collected. In addition, we collect cases that are infrequent or statistically rare. For example, all vowels with the feature pharyngealized are collected, because the feature is rare and overlaps with the feature "retracted tongue root." Similarly, all cases of three-way contrasts in any feature dimension are collected, because they are far less frequent than two-way contrasts.

When interpreting exceptions, we follow two principles given above: the Principle of Contrast and Known Feature First. For illustration, let us consider vowels in Woisika, where we found a three-way contrast in backness. According to UPSID, Woisika has the vowels in Table 2.3, where the diacritic marks [_] and [+] on [i e u o] are not contrastive and can be ignored.

The low vowels [a a̠ ɑ] seem to form a backness triplet. However, the original source, Stokhof (1979: 59), gives the vowel inventory in Table 2.4.

In Stokhof's analysis, there is no three-way contrast in backness among low vowels. In particular, "/ae/ and /ao/ are tension irrelevant", while /â/ and /a/ differ in tenseness. Therefore, three binary features "round," "back," and "tense" suffice to

Table 2.3 Vowel inventory in Woisika, as given in UPSID

	Front	Central	Back
High	ɪ, i_		ʊ, u+
Mid	ɛ, e_		ɔ, o+
Low	a	a̱	ɑ, ɒ

Table 2.4 Vowel inventory in Woisika, as given in Stokhof (1979: 59). Tense vowels are indicated by [ˆ]

	front		central		back	
	lax	tense	lax	tense	lax	tense
high	i	î			u	û
mid	e	ê			o	ô
low	ae		a	â	ao	
	unrounded				rounded	

Table 2.5 Reanalysis of low vowels in Woisika, where the value "central" is not needed for backness

Stokhof	ae	â	a	ao
IPA	æ	a	ɑ	ɒ
Round	–	–	–	+
Back	–	+	+	+
Tense		+	–	

distinguish the low vowels. This is shown in Table 2.5, where IPA transcription is added. We also follow Stokhof and leave open the tense value for /ae/ and /ao/.

The analysis follows from Known Feature First: Since [tense] is already used in this language (possibly related to "advanced tongue root" in other languages), we should prefer this feature to a new feature value, "central." The analysis shows that Woisika is not a compelling case for a three-way contrast in backness. It is possible that another language will be a compelling case. It is also possible that no language is. In the following chapters we shall examine every phonetic dimension in UPSID and P-base, and check every exceptional case.

It may appear that the reinterpretation of exceptional cases is very easy to do, and one may wonder if the process is too liberal. There are, however, strict conditions on the reinterpretation though. First, the reinterpretation observes the Principle of

Contrast. Second, it preserves the original relationship among the sounds. For example, if two sounds A and B contrast in tongue height where A is higher than B, the reinterpretation must preserve the relation. Third, the reinterpretation is limited to one degree of change in the same phonetic dimension. For example, [æ] can be reinterpreted as [ɛ] (one degree higher) but not [e] (two degrees higher). Thus, reinterpretation is highly restricted. There is no guarantee this can always be done. When it appears to be easy to do, it is likely because the original transcriptions are not compelling and reasonable alternative analyses are available.

2.6 Summary

I have introduced the data and the method of the present study. Our method follows three principles, the Principle of Contrast, Maxima First, and Known Feature First, given in (1)–(3). The Principle of Contrast is fundamental to phonology and is not controversial; it determines which phonetic properties need to be represented by features and which not. Maxima First states that, to offer a feature analysis in any language, we must first examine all languages and determine the maximal number of possible contrasts in every phonetic dimension. Known Feature First minimizes theoretical redundancy, so that if a contrastive property can be represented by a feature known to be needed, there is no need to create a new representation (a new feature, feature value, or feature combination) for it.

I have also shown that Maxima First and Known Feature First follow from the Principle of Contrast: To determine whether two sounds contrast and what their contrastive difference is, we must first determine what features there are and the maximal number of contrasts in each feature (hence Maxima First). In addition, if a contrastive difference between two sounds A and B can be represented by a known feature and so there is no extra contrast between A and B, there is no basis for introducing a new feature for them (hence Known Feature First).

Finally, I have illustrated how the proposed method works. Exceptions are defined broadly, a wide net is cast, and anything that seems controversial is identified and collected exhaustively. All exceptions are then examined individually, under the Principle of Contrast. We now move on to chapters that report the actual analyses of the data and the results.

3

Vowel contrasts

This chapter analyzes vowel contrasts in UPSID and P-base. I begin with a detailed examination of vowels in UPSID. Then I examine vowels in P-base, abbreviating cases that are similar to those in UPSID and focusing on additional contrasts not reported in UPSID.

3.1 Vowels in UPSID

Following Maddieson (2005), I divide vowels into a "basic" and a "non-basic" set. Basic vowels are monophthongs that involve tongue positions and lip rounding only. Non-basic vowels can be divided into further categories based on what additional properties they have, such as nasality, length, breathiness, or diphthongs. Table 3.1 shows vowel categories in UPSID, where a "type" refers to a distinct transcription. If a vowel falls under two (or more) categories, it is arbitrarily included in just one. For example, nasalized diphthongs are included in "diphthong" and not in "nasalized"; this is because every category is accounted for, so it does not matter which category nasalized diphthongs are grouped with. A token is an occurrence of a transcription. A type is a distinct transcription, regardless of how many times it is found.

I shall examine all cases that involve uncommon or controversial contrasts. For example, "pharyngeal" and "extra-short" vowels are uncommon, which will be examined. In addition, some linguists believe that no feature needs a three-way contrast (e.g. Jakobson et al. 1952; Chomsky and Halle 1968); therefore, all cases of three-way contrast will be examined.

I assume that diphthongs count as two vowels each, and that long vowels are linked to two timing slots or moras (Chapter 1); therefore, they require no further discussion. In addition, nasalization, creakiness (called "laryngeal" in UPSID), and breathiness are well-known and non-controversial features that can be added to other vowels; therefore, they too require no further discussion. We shall examine all other categories: "basic," "extra-short," "pharyngeal," "voiceless," "retroflex," and "fricative" vowels.

A Theory of Phonological Features. First edition. San Duanmu.

Table 3.1 Vowels in UPSID. The category "laryngeal" is also called "creaky" in the literature. Categories with * are examined below

Category	Type	Token
Basic*	45	2,699
Diphthong, long	129	488
Laryngeal, breathy	24	68
Nasalized	30	508
Extra-short (overshort)*	19	29
Pharyngeal*	11	23
Voiceless*, retroflex*, fricative *	11	18
All	269	3,833

3.1.1 Basic vowels

Features for basic vowels in UPSID are similar to those in its predecessor, the Stanford Phonology Archive (Crothers et al. 1979). They include (i) three degrees of backness, (ii) two degrees of height for high vowels, (iii) three degrees of height for mid vowels, and (iv) two degrees of height for low vowels. The system yields a table of seven rows and six columns, shown in Table 3.2. We examine the controversial aspects of this system, which are properties (i), (iii), and (iv).

Seven of the basic vowels in UPSID have an additional diacritic symbol and do not fit into the table. An examination shows that none of them contrasts with a regular vowel. The result is given in Table 3.3.

In each case, the vowel with a diacritic does not contrast with one without. For example, [e̝_] is found in Khalkha and Karen; Khalkha does not have [e] and Karen does not have [ə]. Following the Principle of Contrast, the special symbols can be replaced by regular ones. In fact, some symbols in our analysis are used by other authors. For example, Stokhof (1979: 59), the source for Woisika, uses [e o] rather

Table 3.2 38 basic vowels in UPSID. Vowels in parentheses are not found

	Front		Central		Back	
High	i	y	ɨ	ʉ	ɯ	u
High (lower)	ɪ	ʏ	ɪ	ʊ	Ɯ	ʊ
Mid (higher)	e̝	ø̝	ə̝	ɵ̝	ɤ̝	o̝
Mid	e	ø	ə	ɵ	ɤ	o
Mid (lower)	ɛ	œ	ɜ	ɞ	ʌ	ɔ
Low (raised)	æ	(æʷ)	ɐ	(ɐʷ)	ɐ̠	ɐ̠ʷ
Low	a	(aʷ)	a̠	(a̠ʷ)	ɑ	ɒ

Table 3.3 Seven "basic" vowels that have an additional diacritic, none of which contrasts with a more common vowel without a diacritic

Vowel	Description	Languages	Inventory	Analysis
e̝_	Retracted [e̝]	Khalkha, Karen	No [e] or [ə]	[e] or [ə]
e_	Retracted [e]	Woisika	No [e]	[e]
ø̝_	Retracted [ø̝]	Malakmalak	No [ø]	[ø]
ɤ̝+	Fronted [ɤ̝]	Mandarin, Hmong	No [ɤ]	[ɤ]
o̝+	Fronted [o̝]	Khalkha	No [o]	[o]
o+	Fronted [o]	Woisika	No [o]	[o]
i-	Velarized [i]	Siriono	No [ɯ]	[ɯ̺]

than [e_ o+]; Jones (1962: 8), the source for Karen, uses [ə], rather than [e̝_]; and for Mandarin, Duanmu (2007) uses [ɤ] instead of [ɤ̝+].

Next we consider the interpretation of [i-] as [ɯ̺] (an apical [ɯ]). The source, Priest (1968: 103), transcribes the vowel as [ɨ] but describes it as "high, close, front, unrounded vowel with friction and back tongue rounding or grooving". If backness is a property of the tongue body, there is a problem with the description: its "front" part requires the tongue body to be front, while its "back" part requires the tongue body to be back, but it is impossible for the tongue body to be both front and back at the same time. If [ɨ] is [ɯ̺], there is no problem. In [ɯ̺], the tongue body is indeed back (for [ɯ]), while the tongue tip is front. Since the gestures involve different articulators, they are compatible. Alternatively, the sound could simply be [ɯ]: In Priest's own transcription, Siriono has five oral vowels [i ɨ u e o a] (plus their nasal counterparts), where there is no contrast between [ɨ] and [ɯ].

Let us now consider three-way contrast in backness and height. For backness, there are twelve triplets, shown in (1).

(1) Triplets for three-way contrast in backness
[i ɨ ɯ], [ɪ ɪ Ɯ], [e̝ ə̝ ɤ̝], [e ə ɤ], [ɛ ɜ ʌ], [æ ɐ ɐ̝], [a a̠ ɑ]
[y ʉ u], [ʏ ʊ ʊ], [ø̝ ɵ̝ o̝], [ø ɵ o], [œ ɞ ɔ]

We search through the 451 inventories in UPSID for each of the triplets. Three are found, each in one language, shown in (2).

(2) Backness triplets found in UPSID

Triplet	*Found in*
[e ə ɤ]	Moro
[i ɨ ɯ]	Nimboran
[a a̠ ɑ]	Woisika

We have seen in Chapter 2 that Woisika does not need a backness triplet. The case in Moro is weak, too. According to UPSID, Moro has seven vowels [i u e ə ɤ o a].

However, according to the source, Black and Black (1971), [ə] is not a full phoneme: It is found in unstressed syllables only. Therefore, [e ɤ] are the only vowel phonemes that are mid and unrounded, and there is no three-way contrast.

Finally, consider Nimboran, which has six vowels [i ɨ ɯ e ɤ a], all of which are unrounded (Anceaux 1965: 9). The backness triplet is supposed to be [i ɨ ɯ]. However, according to the source (Anceaux 1965: 13–15), for some speakers [ɨ] is "rather tense" and "backed," whereas [ɯ] is slightly lowered. This means that [ɨ ɯ] could differ in tenseness, while both being back and high.

The discussion shows that there is no compelling evidence for a three-way contrast in backness. Next, we consider three-way contrast in height among mid vowels. There are six such triplets, shown in (3).

(3) Six contrastive triplets in height among mid vowels
[e̝ e ɛ], [ø̝ ø œ], [ə̝ ə ɜ], [ɵ̝ ɵ ɞ], [ɤ̝ ɤ ʌ], [o̝ o ɔ]

An exhaustive search through UPSID yields just two of the triplets, shown in (4), both from the language Klao.

(4) Height triplets for mid vowels found in UPSID

Triplet	Found in
[e̝ e ɛ]	Klao
[o̝ o ɔ]	Klao

According to UPSID, Klao has nine oral vowels, all of which except [e̝ o̝] have a nasal counterpart. The oral vowels are shown in (5).

(5) Oral vowels in Klao, as given in UPSID

High	i		u
Mid (higher)	e̝		o̝
Mid	e		o
Mid (lower)	ɛ		ɔ
Low		a	

However, in the source, Singler (1979: 63), the oral vowels are transcribed as in (6), where [e o] have "expanded pharynx" and [ẹ ọ] have non-expanded pharynx.

(6) Oral vowels in Klao, as given in the source (Singler 1979: 63)

	Non-back	Back
High	i	u
Mid	e ẹ	o ọ
Low	ɛ	a ɔ

According to Singler (1979), [ɛ ɔ] are not mid but low vowels. Therefore, there is no three-way contrast in height among mid vowels.

Before we move on to another feature, let us also examine the height contrast among low vowels, which is less common than that among high or mid vowels. In Table 3.2, there are four pairs of low vowels that contrast in height, repeated in (7).

(7) Transcribed low vowel pairs that contrast in height
[æ a], [ɐ a̠], [ɐ̠ ɑ], [ɐ̠ʷ ɒ]

A search through UPSID shows that only the pair [ɐ a̠] is found. Only seven languages contain it, shown in (8).

(8) Languages in UPSID that contain [ɐ a̠]
Amo, Dizi, Kam, Mazahua, Nunggubuyu, Ogbia, and Vietnamese

Let us examine each case. As before, following the Principle of Contrast, we ask whether the extra contrast is warranted. Some possibilities are shown in (9).

(9) Alternative analyses of [ɐ a̠]

Analysis	*Example*	*Comment*
[ʌ a]	Amo	[a] is low but [ʌ] is mid.
[a ɑ]	Kam	The contrast is in backness (not in height).
[a aː]/[ɑ ɑː]	Nunggubuyu	The contrast is in length (not in height).

According to Di Luzio (1972: 5), a source of UPSID, Amo has six vowels, shown in (10), where only one is low. Both the features and the phonetic symbols are in the original.

(10) Vowel phonemes in Amo, as given in Di Luzio (1972: 5)

	High		Center		Back
High	i				u
Low–mid		e	ö	o	
Low			a		

However, UPSID lists ten vowels for Amo, shown in (11), "on the basis of field work by L. Hyman and analysis of tape recordings."

(11) Vowel phonemes in Amo, as given in UPSID

	High	Central	Back
High	i		u
	ɪ		ʊ
Mid	e̞		o̞
	ɛ		ɔ
Low		ɐ	
		a̠	

Still, it can be seen that there is no mid vowel [ʌ], [ɤ], or [ə]. Therefore, we can represent the contrast between [ɐ] and [a] as [–low] and [+low], instead of [+low, +raised] and [+low, –raised].

Dizi has six vowel phonemes, given in UPSID as [i e a̠ ɐ o u]. Again, there is no [ʌ], and we can consider [ɐ] to be [ʌ]. Indeed, this is the analysis of Allen (1976), the source of UPSID, shown in (12), where there is just one low vowel.

(12) Vowel phonemes in Dizi, as given in Allen (1976: 378)

High	[i u]
Mid	[e ʌ o]
Low	[a]

Kam has seven vowels, given in UPSID as [i e ɐ a̠ ə o u]. However, Yang (2009) transcribes [ɐ a̠] as [a ɑ] instead, which differ not in height but in backness. There is some evidence for Yang's analysis. Consider vowel alternation conditioned by tones, shown in (13). Tones are orthographically marked with a consonant at the end of a syllable (-s, -t, -x, -l, -p, -c, etc.). The pitch values of the tones need not concern us.

(13) Vowel alternation conditioned by tone, when the coda is [p], [t], or [k]

Tones	-s -t -x	-l -p -c
Mid vowel	[e]	[ə]
Low vowel	[a]	[ɑ]

We see an alternation between the mid vowels [e] and [ə], where [e] occurs with the tones -s, -t, and -x and [ə] occurs with the tones -s, -t, and -x. Similarly, there is an alternation between the low vowels [a] and [ɑ]. The [e]–[ə] difference is one of backness. Therefore, it would be natural if the [a]–[ɑ] difference is one of backness, too.

Mazahua has nine oral vowels and six nasal vowels. In UPSID they are transcribed as [i e̞ ɛ a̠ ɐ ə ɔ o̞ u] and [ĩ ẽ̞ ã̠ ə̃ õ̞ ũ]. However, the source, Spotts (1953: 254), transcribes the oral vowels as in (14), where [ʌ] is mid (changed to [ɐ] in UPSID) and [a] is the only low vowel.

(14) Oral vowels in Mazahua, as given in Spotts (1953: 254)

i		u
e	ə	o
ɛ	ʌ	ɔ
	a	

Nunggubuyu has four vowels, given as [i a ɐ ɯ] in UPSID, all being unrounded. However, in one of the sources, Hughes and Leeding (1971), the vowel inventory is given as in (15), where [æ] is in parentheses for being "rare."

(15) Vowels in Nunggubuyu, as given in Hughes and Leeding (1971: 72)

	Front –round	Central –round	Back –round
High close			
High open	i		u
Low close	(æ)	a	
Low open		a:	

Note that [a a:] differ in length. As Hughes and Leeding (1971: 75) point out, [a:] is not only long but is the only vowel that always attracts stress. If we assume a length contrast, there is no need to assume a height contrast for [a a:]. In fact, Heath (1984), another source for Nunggubuyu, argues that not only can [a] be long, but [i u] can as well. In other words, Nunggubuyu has six vowels, [i a u i: a: u:] (Heath 1984: 9). However, Heath (p. 9) also points out that "length contrasts are of very little functional interest," because "only a handful" of minimal pairs can be found.

Ogbia has ten vowels, given in UPSID as [i e̞ a o̞ u ɪ ɛ ɐ ɔ ʊ]. However, the source, Williamson (1972: 1), transcribes them as [i e a o u ɪ ɛ ʌ ɔ ʊ], where [a] is the only low vowel. There seems to be no need to revise the original transcriptions.

Vietnamese has fourteen vowels, three of which are diphthongs. The monophthongs are given in UPSID as [ɪ e̞ ɛ a̠ ɐ ʌ ɔ ɤ̞ o̞ ɯ ʊ]. In the source, Thompson (1965: 20), they are given as in (16). The feature labels are as in the original. I have converted the transcription to IPA but kept the original ones for the three low vowels.

(16) Vietnamese monophthongs, as given in Thompson (1965: 20)

	Front –round	Central –round	Back +round
High	i	ɯ	u
Mid (upper)	e	ɤ	o
Mid (lower)	ɛ		ɔ
Low (upper)	â	ă	
Low (lower)	a		

UPSID has converted Thompson's "upper low" [ă] to the lower mid back vowel [ʌ]. This seems reasonable, since Thompson says that it is similar to the vowel in the English word *but*. UPSID has also converted Thompson's front low vowels [â a] to central [ɐ a̠], which need not concern us. What is relevant is that, as Thompson points out, there is a length difference between [â a], where [â] is "very short" and only occurs in a closed syllable (one that ends in a consonant). Therefore, an alternative is to interpret the contrast between [â a] (or between [ɐ a̠] in the transcription of UPSID) as one of length, rather than one of height. In this regard, it is interesting to consider yet another analysis, that of Ngo (1999), shown in (17), which makes use of a length contrast for two vowel pairs. The two short vowels are mentioned by Thompson (1965) as well.

(17) Vietnamese monophthongs, as given in Ngo (2001: 9)

	Front	Central	Back
High	i	ɯ	u
Mid	e	ə əː	o
Low	ɛ	a aː	ɔ

A comparison of the three analyses is shown in (18), where we focus on four vowels that are of most relevance.

(18) Different analyses of three Vietnamese vowels

Thompson (1965)	ɤ	ă	â	a
UPSID	ɤ̞	ʌ	ɐ	a̠
Ngo (2001)	əː	ə	a	aː

In summary, in UPSID, there is no compelling evidence of three-way contrast in any vowel feature, if we treat high, mid, and low as separate features. In addition, evidence for a height contrast among low vowels is inconclusive, since alternative analyses are available. Thus, the inventory of basic vowels in UPSID can be reduced from Table 3.2 to Table 3.4, where the vowel [Œ] is not found in UPSID.

It can be seen from Table 3.4 that the vertical axis offers five degrees of contrast. This dimension remains odd in feature theory, since there is very little evidence that we need more than a two-way contrast in any other dimension; yet all previous studies agree that we need no fewer than five degrees of contrasts here. Some linguists have attempted to address the problem by decomposing the vertical axis in Table 3.4 into three features, such as [high], [low], and [tense] (Chomsky and Halle 1968), as shown in (19).

(19) Decomposing multiple contrasts into multiple binary features

UPSID term	Binary features
High	[+high, –low, +tense]
High (lower)	[+high, –low, -tense]
Mid	[–high, –low, +tense]
Mid (lower)	[–high, –low, –tense]
Low	[–high, +low]

TABLE 3.4 Basic vowels in a contrast-based analysis of UPSID

	Front		Back	
High	i	y	ɯ	u
High (lower)	ɪ	ʏ	Ɯ	ʊ
Mid	e	ø	ɤ	o
Mid (lower)	ɛ	œ	ʌ	ɔ
Low	æ	(Œ)	ɑ	ɒ

Chomsky and Halle (1968) have offered some argument for the decomposition, but the proposal is not always accepted. In addition, the lack of contrast in [tense] among low vowels remains to be explained, as does the feature [tense] itself. I shall return to this issue in Chapter 4 when we discuss vowel height.

3.1.2 Extra-short (overshort) vowels

The presence of extra-short vowels (termed "overshort" vowels in UPSID) calls for a three-way contrast in length: extra-short, regular, and long. It requires either a new value for length or a new structural representation of time.

An often-cited case of three-way contrast in length is Estonian (Lehiste 1960), but Prince (1980) argues that, if prosodic factors are taken into consideration, Estonian only has a two-way length contrast. Because of the controversy, we examined UPSID and looked for length triplets in every vowel. As shown in (20), the search yielded no hit.

(20) Search for length triplets in UPSID
Length triplets: [iː i ĭ], [aː a ă], etc.
Triplets found: None

Nevertheless, UPSID contains some transcriptions of "extra-short" vowels, found in seven languages, totaling twenty-nine tokens. The data are summarized in Table 3.5.

Since none of the languages has a three-way contrast in length, why is the length difference represented as short (regular) vs extra-short, instead of long vs regular? It is not the case that UPSID never uses long vs regular for a length difference. For example, vowels in Bardi are [i a ɔ u iː aː uː] and those in Totonac are [i æ u iː aː uː], where the length difference is represented as long vs short.

An examination of the languages shows that extra-short vowels are used in four situations, summarized in Table 3.6, where "flex" refers to vowels that can be long or short.

TABLE 3.5 Languages that contain extra-short vowels in UPSID

Language	Extra-short vowels	Count
Po-Ai	[ă ɨ̆ ɪ̆ ɔ̆ ŭ ɛ̆ ŏ̞]	7
Sebei	[ə̆ ɵ̆ ɨ̆ ɪ̆ ɐ̆ʷ ɐ̆ ʉ̆]	6
Khanty	[ɵ̆ ĕ ŏ ɒ̆]	4
Lungchow	[ă ɪ̆ ʊ̆ ɯ̆]	4
Nenets	[ă ɪ̆ ŭ]	3
Chuvash	[ɔ̆ ø̞̆]	2
Angaatiha	[ə̆]	1
Chukchi	[ə̆]	1
Georgian	[ə̆]	1

TABLE 3.6 Situations where extra-short vowels are used in UPSID

Situation	Transcription	Languages
Short vs long	Extra-short vs regular	Sebei
Flex vs long	Extra-short vs regular	Po-Ai, Lungchow
Short vs flex	Extra-short vs regular	Khanty, Nenets
Short vs regular (with no contrast)	Extra-short vs regular	Chuvash, Chukchi, Angaatiha, Georgian

In none of the cases is it necessary to use extra-short vowels. In particular, a long vowel is linked to two timing slots (Chapter 1). A vowel with flexible length means that it is linked to one timing slot in some cases and two in others, rather than always being extra-short (as in Po-Ai) or always being regular (as in Khanty). Finally, when there is no contrast in length, there is no need to use a length feature, even if some vowels are phonetically shorter than others (as in Chuvash). Let us look at an example of each case.

Sebei has twelve vowels, which form six long–short pairs. They are given in UPSID as [i e̞ a ɔ o̞ u ɨ̆ ə̆ ɐ̆ ɐ̆ʷ ɵ̆ ʉ̆]. Montgomery (1970), the source, represents the vowels in (21), where short vowels are centralized.

(21) Vowel phonemes in Sebei, as given in Montgomery (1970: 50)

	Front	Central	Back
High	i		u
		ï ü	
Mid	e	ë ö	o
		ä ɔ̈	
Low	a		ɔ

Since there is only a two-way contrast in length, we can represent the difference as long vs short, rather than short vs extra-short. Alternatively, there could be a stress difference, since Montgomery (1970: 51) points out that "the distribution of short vowels is limited to word-medial occurrences," "/ä/, /ö/ and /ɔ̈/ occur rarely," and "contrasts within the set of short vowels have a low functional yield." In other words, there is a possibility that the short vowels are unstressed versions of other vowels. Unfortunately, the focus of Montgomery (1970) was on orthography and there was no discussion of stress in the language.

According to UPSID, Po-Ai has sixteen vowels, given as [i ɨ u e̞ ə̞ ɛ a ɔ ĭ ɨ̆ ŭ ŏ̞ ɛ̆ ă ɔ̆ aɪ], where [ĭ ɨ̆ ŭ ŏ̞ ɛ̆ ă ɔ̆] are "extra-short." According to Li (1977: 13), the source of UPSID, Po-Ai has just nine vowel phonemes, given in (22), where "[t]he high and low vowels may be either long or short, but the mid vowels are always long."

(22) Vowel phonemes in Po-Ai, as given in Li (1977: 13)

	Front	Central	Back
High	i	ï	u
Mid	e	ə	o
Low	ɛ	a	ɔ

The UPSID list seems erroneous in several ways. First, with regard to length, Li only mentions a difference between long and short; there is no need to interpret it as one between short and extra-short. Second, according to Li, all mid vowels are long. Therefore, the UPSID [ŏ̞] should be [o̞] instead. Third, Po-Ai has several diphthongs, which Li considers to be vowel clusters, instead of additional phonemes. It is unclear why UPSID lists just one of the diphthongs [aɪ], instead of listing all, or none. Finally, since vowel length can be represented by timing slots (see Chapter 1), there is no need to consider long vowels to be additional phonemes. In this regard, Li's original analysis seems better than the UPSID reinterpretation.

Khanty has thirteen vowels, given in UPSID as [i ʉ ɯ u e ɵ o æ a ĕ ɵ̆ ɒ̆ ŏ]. The source, Gulya (1966: 23), represents them in (23), where 9 of them are "full" and 4 "reduced."

(23) Vowel phonemes in Khanty, as given in Gulya (1966: 23)

	Palatal		Velar	
	Non-labial	Labial	Non-labial	Labial
Close	i	ü	i̲	u
Middle	e	ö		o
Open	ä		a	
"Reduced"	ə	ö̆	ă	ŏ

If we interpret "palatal" as front and "velar" as back, then UPSID has misrepresented three of the vowels: the palatal labials [ü ö ö̆] should not be central [ʉ ɵ ɵ̆] but front [y ø ø̆] instead. With regard to length, Gulya (1966: 24) states that the main difference between "full" and "reduced" vowels is that the former can be long or short (depending on context, ranging from 100ms to over 200ms), whereas the latter are always short (80ms or so). In addition, all vowels can carry stress, although full vowels tend to attract it (Gulya 1966: 36–7). This means that "reduced" vowels are not extra-short or reduced (i.e. unstressed), but of regular length. Given this, a viable analysis of Khanty vowels is in (24).

(24) Analysis of Khanty vowels

	Front	Back
High	i y	ɯ u
Mid	e ö	o
	ɛ œ	ʌ ɔ
Low	æ	ɑ

The "reduced" vowels can be seen as [−ATR] mid vowels, which are short, whereas other vowels are [+ATR] and can be short or long. There is no need to assume extra-short vowels.

Angaatiha has two diphthongs [ai au] and seven monophthongs. The monophthongs, as given in UPSID, are shown in (25).

(25) Monophthongs in Angaatiha, as given in UPSID

Long	[a:]
Short	[i u e o a]
Extra-short	[ə̆]

We can represent [a:] with two timing slots. For other vowels there is no length contrast, because [ə̆] can be distinguished from others by regular features. For example, Huisman (1973: 43) and Huisman et al. (1981: 52), two sources of UPSID, represent the vowel phonemes in (26), where [ai au] are not included and [ə] can be distinguished from others by the features "high," "non-round," and "non-front."

(26) Vowel phonemes in Angaatiha, as given in Huisman (1973: 43)

	Front	Central	Back
High	i	ə	u
Low	e	a a:	o

There are other possible analyses that we shall only mention briefly here. For example, Huisman et al. (1981: 57) note that [ə] can alternate with [a o e], which suggests that [ə] may be a reduced version of them (and hence not a full phoneme itself). Similarly, Huisman et al. (p. 56) note that [i u] are shorter than [a], which in turn is shorter than [a:], in agreement with a well-known phonetic effect that lower vowels are longer than higher vowels. In addition, Huisman et al. (p. 56) note that [a] can alternate with [ʌ]. This suggests that we could represent the contrast between [a]-[a:] as [ʌ]–[a] instead, i.e. one of height, rather than one of length.

Let us consider one more language, Georgian. UPSID lists six vowels for it [ɪ e a ə̆ o ʊ]. Aronson (1997) lists just five [i ɛ ɑ ɔ u]. Robins and Waterson (1952), the source of UPSID, also lists five vowels [i e a o u]. Robins and Waterson (p. 66) point out that sometimes voiced initial or final stops "are followed by a distinct voiced "off-glide," e.g. /kargad/ [kˀaɹgad̪ᵊ] 'well (adverb)', /gdeba/ [g̊ᵊd̪eba] 'to throw'. It is clear, though, that this epenthetic sound has no contrastive function and is not a full phoneme.

In summary, there is no compelling evidence for extra-short vowels in UPSID, and there is no basis for creating a new feature for them.

3.1.3 Pharyngeal vowels

Pharyngeal vowels are made with a narrowed pharynx, achieved by retracting the tongue root. The gesture itself is not unusual, but there is a more common term for the action, namely, "advanced tongue root" [ATR] (Jakobson and Halle 1964;

Table 3.7 Languages that contain pharyngeal vowels in UPSID

Language	Pharyngeal vowel	Contrast with [–ATR] V?
Archi	Yes	No
Even	Yes	No
Hamer	Yes	No
Lak	Yes	No
Neo-Aramaic	Yes	No
!Xu	Yes	No

Ladefoged 1964; Stewart 1967; Halle and Stevens 1969; Perkell 1969, 1971; Wood 1982; Lindau 1979; Vaux 1996; Kenstowicz 2009; Allen et al. 2013). The question then is: are pharyngeal vowels different from [–ATR] vowels? To find out, we examine every language that has pharyngeal vowels and see whether they contrast with [–ATR] vowels, such as [ɪ ɛ ʌ ɔ ʊ], which have also been called "lax" vowels.

A search through UPSID yields six languages that contain pharyngeal vowels. The data are summarized in Table 3.7. None of the vowels contrasts with [–ATR] or lax vowels.

In Archi, Lak, Neo-Aramaic, and !Xu, no lax vowel symbol is used. Therefore, we can interpret pharyngeal vowels as lax vowels. This is illustrated with Archi in (27).

(27) Analysis of pharyngeal vowels in Archi

UPSID	i	e	a	o	u	iˤ	eˤ	aˤ	oˤ	uˤ
Analysis	i	e	a	o	u	ɪ	ɛ	ʌ	ɔ	ʊ

In Even, the pharyngeal diacritic is redundant, since it overlaps with lax vowels. Therefore, we can omit the pharyngeal diacritic, shown in (28).

(28) Analysis of pharyngeal vowels in Even

UPSID	ɪ	ɛ	ɑ	o	u	ɪˤ	ɔˤ	ʊˤ	ie	ia
Analysis	i	e	ɑ	o	u	ɪ	ɔ	ʊ	ie	ia

The case in Hamer is similar to that in Even, where the pharyngeal diacritic is again redundant. This is shown in (29).

(29) Analysis of pharyngeal vowels in Hamer

UPSID	i	e	a	o	u	ɪˤ	eˤ	ʌˤ	ɔˤ	ʊˤ
Analysis	i	e	a	o	u	ɪ	ɛ	ʌ	ɔ	ʊ

In conclusion, pharyngeal vowels overlap with lax vowels, and no language in UPSID has a contrast between the two. There is therefore no need to keep both terms.

3.1.4 *Other vowels*

There are three more categories to consider: voiceless vowels, retroflex vowels, and fricative vowels. There is no question that voiceless vowels can be pronounced. In whispered speech, for example, all vowels (and consonants) are voiceless. What remains to be seen is whether any language uses voiced and voiceless vowels, such as [a] vs [ḁ], contrastively, rather than as different styles of speech that are non-contrastive. To find out, we first search for all voiceless vowels in UPSID. The result is shown in Table 3.8.

Next we ask whether the voiceless vowels are productive, or whether they only occur in restricted contexts. The result is summarized in Table 3.9.

In all the cases, voiceless vowel phonemes are found in restricted contexts. In fact, original sources for all the languages consider them to be allophones of regular vowels, rather than independent phonemes. In particular, Ray (1967: 10), the source on Dafla, gives the vowel inventory as [i e ɑ ʌ o ɯ u]. Tucker et al. (1977: 301), a source on Sandawe, gives the phoneme inventory as [i e a o u]. Tucker (1971: 341), a source on Ik, gives the phoneme inventory as [i ɪ e ɛ a ɔ o u ʊ]. No source contains voiceless vowels. It seems that the compilers of UPSID either made clerical errors, or were more generous than the original authors in entertaining voiceless vowels.

Next we consider retroflex vowels. UPSID lists three, totaling five tokens, found in four languages. The data are summarized in Table 3.10.

Table 3.8 Voiceless vowels in UPSID

Vowel	[i̥]	[u̥]	[ḁ]	[e̥]	[o̥]
Found in	Dafla, Ik, Sandawe	Dafla, Ik, Sandawe	Ik	Ik	Ik

Table 3.9 Contexts where voiceless vowels are found

Language	Context
Sandawe	Word-final position after a limited set of consonants; almost always low-toned
Ik	Word-final position
Dafla	Word-final position

Table 3.10 Retroflex vowels in UPSID

	[ɚ]	[a˞]	[ɨ˞]
Found in	Gelao, Mandarin, Naxi	Mandarin	Tarascan

A common (but possibly not the only) way to produce a retroflex vowel is to curl the tongue tip up and towards the post-alveolar region. This articulation is easier for back vowels, where there is more room in the front of the mouth for the tongue tip gesture. In this regard, it is relevant to note that all three retroflex vowels are non-front. In fact, more cases of retroflex vowels can be found. For example, in American English, the rime in *fur* can be seen as a retroflex vowel [ɚ] or [ɝ]. Similarly, Standard Chinese has at least [u˞] and [o˞], besides [a˞] (or [ɑ˞]) and [ɚ], although these vowels can be seen as marginal phonemes, resulting from a combination of a regular vowel plus a diminutive suffix that consists of a retroflex feature (Duanmu 2007).

In summary, it is possible to pronounce various retroflex vowels (especially when the vowel is non-front), and no new feature is involved. However, it is interesting to note how rarely they occur as contrastive phonemes.

Finally, let us consider fricative vowels, which are like fricatives phonetically, having at least some frication, but vowels phonologically, serving as the rime or nucleus of a syllable. This definition is not precise, though. For example, syllabic consonants, such as [z] and [ʐ] in Standard Chinese (Chao 1968; Duanmu 2007) and [z ʒ ʒʷ β] in Nantong Chinese (Ao 1993), have been called vowels (Karlgren 1915–26) or fricative vowels (Ao 1993), even though they are not independent phonemes. Similarly, a vowel may sometimes be pronounced with some degree of frication, such as [fu] 'rich' in Standard Chinese, which can be realized as [fʋ] or [fv]; and such pronunciations can be called fricative vowels, too. In neither case is the fricative vowel an independent phoneme, since there is no contrast between it and a syllabic fricative, or between it and a regular vowel.

To establish a true case for fricative vowels, we need to find two kinds of contrast: (a) a fricative vowel vs a syllabic fricative, and (b) a fricative vowel vs a regular vowel. UPSID lists three fricative vowel types, totaling four tokens, found in three languages. The data are summarized in Table 3.11, where I denote a fricative vowel with [^{F}].

In Ewondo, [i u] (and especially [i]) are often accompanied by some frication (Redden 1979: 8), but there is no contrast between a fricative vowel and a regular vowel. Similarly, in Bai and Naxi, the fricative vowel [ʋ^F] is not an independent phoneme, since it does not contrast with a syllabic [v]. For example, in his analysis of Naxi, Jiang (1980: 60) uses the same symbol [v] for both the fricative and its syllabic counterpart [ʋ]. A similar analysis is offered by Xu and Zhao (1984) for Bai.

In conclusion, all fricative vowels in UPSID can be analyzed as either syllabic fricatives or allophones of high vowels. There is no basis for introducing a new vowel feature.

Table 3.11 Fricative vowels in UPSID

Vowel	[ʋ^F]	[i^F]	[u^F]
Tokens	2 (Bai, Naxi)	1 (Ewondo)	1 (Ewondo)

3.2 Vowels in P-base

Having examined all vowels in UPSID, we turn to vowels in P-base. As in the analysis above, we divide P-base vowels into various categories, shown in Table 3.12. As with UPSID, if a vowel falls under two (or more) categories, it is arbitrarily included in just one. For example, long nasal vowels are grouped with "long" and not with "nasal."

As discussed before, long vowels and diphthongs can be analyzed in terms of two timing slots, and need no further discussion. Non-syllabic vowels are reported in three languages: Angami, Bengali, and Romanian. They occur next to a regular vowel in a diphthong, such as [i̯o e̯o] in Angami, where [i̯ e̯] are non-syllabic vowels. The diacritic in [i̯ e̯] indicates that the diphthong is a single syllable and that the non-syllabic vowel is not the nucleus (or the peak) of the syllable. However, if syllable structure is independently provided, it is clear which part of a diphthong is in the nucleus, and there is no need for a separate diacritic. Nasal and breathy features are also well known and non-controversial, as discussed above, and need no further discussion.

The remaining vowels do need some discussion. In particular, we want to know (i) whether there is any three-way contrast among basic vowels, (ii) whether pharyngeal vowels can contrast with [–ATR] vowels, (iii) whether creaky vowels can contrast with glottal vowels, (iv) whether voiceless vowels can contrast with regular vowels, (v) whether extra-short and extra-long vowels call for a three-way contrast (or more degrees of contrast) in length, and (iv) whether additional vowel types in the category "others" require new features. We shall see that in all cases the answer is no. For the sake of completeness, for each unusual contrast I shall list all the languages that seem to have it and whether each turns out to be a valid case after reexamination. To avoid repetitions, however, I shall offer only abbreviated discussions of the re-examination process, most of which are similar to those for UPSID vowels.

TABLE 3.12 Vowel categories in P-base

Category	Type	Token
Basic	26	3,756
Diphthong, long, non-syllabic	103	993
Nasal, glottal, breathy, creaky	85	469
Pharyngeal	9	12
Voiceless	9	9
Extra-short, extra-long	11	15
Others	28	60
All	271	5,314

Table 3.13 Basic vowels in P-base; those in parentheses are not found. The height distinction within high, mid, or low vowels is labeled as tense vs lax

		Front		Central		Back	
High	Tense	i	y	ɨ	ʉ	ɯ	u
	Lax	ɪ	ʏ	(ᵻ)	(ᵿ)	(ɯ)	ʊ
Mid	Tense	e	ø	ə	ɵ	ɤ	o
	Lax	ɛ	œ	ɜ	(ɞ)	ʌ	ɔ
Low	Tense	æ	(æʷ)	ɐ	(ɐʷ)	(ɐ̱)	(ɐ̱ʷ)
	Lax	a	(ɶ)	a̱	(a̱ʷ)	ɑ	ɒ

3.2.1 Basic vowels

P-base assumes a system of 36 possible basic vowels. If we exclude vowels with special diacritics (which are discussed separately), only 26 basic vowels are found. This is shown in Table 3.13, where non-occurring vowels are shown in parentheses. P-base does not offer a specific interpretation of the IPA symbols used. I have added common feature labels (high, mid, low, front, central, back, and tense, lax), which preserve the relative tongue positions among the IPA symbols.

As with UPSID vowels, we focus on two issues: three-way contrast in backness and height contrast among low vowels. In Table 3.13 there are six triplets in backness. For completeness, we also add four triplets of long vowels, since tense vowels are sometimes written as long. The list is shown in (30).

(30) Triplets for three-way contrast in backness
[i ɨ ɯ], [e ə ɤ], [ɛ ɜ ʌ], [a a̱ ɑ], [y ʉ u], [ø ɵ o]
[iː ɨː ɯː], [eː əː ɤː], [yː ʉː uː], [øː ɵː oː]

A search through the P-base inventories yields three triplets, found in four languages. The result is shown in (31).

(31) Backness triplets found in P-base

Triplet	Found in	Valid
[i ɨ ɯ]	Ashuku, Tepecano	No
[y ʉ u]	Saami	No
[ɛ ɜ ʌ]	Mixe (Coatlán variety)	No

A re-examination of the languages shows that none of them demonstrates a compelling case for a three-way contrast in backness; a two-way contrast is sufficient.

Next, we consider height contrast in low vowels. Based on the IPA transcription, there are two such pairs in P-base, [æ a] and [ɐ a̠]. However, since [a] is often used for [a̠] and [ɑ], when we search for [æ a], we added [ɑ] to ensure that [a] is a front vowel. Similarly, we added [ɐ a] in case [a] is used for [a̠] and we added [ɐ ɑ] in case [ɐ] is counted as back. The four pairs of interest are shown in (32).

(32) Low vowel pairs that contrast in height
[æ a (ɑ)], [ɐ a̠], [ɐ a], [ɐ ɑ]

A search through P-base yields three of the pairs, shown in (33). Each of the inventories is then examined and none turns out to be a valid case.

(33) Height contrast among low vowels in P-base

Vowel pair	Found in	Valid
[æ a (ɑ)]	Bengali, German, Gujarati, Maithili	No
[ɐ a]	Portuguese (European and Brazilian)	No
[ɐ ɑ]	Saami	No

The Saami inventory given in P-base contains a clerical error, whereby an extra low vowel is listed. The Portuguese inventory has no [ʌ]; if we interpret [ɐ] as [ʌ], [a] would be the only low vowel. Similarly, German does not have [ʌ] (except [ə], which occurs in unstressed syllables only and is not a full phoneme); if we interpret [ɐ] as [ʌ], there is no [ɐ a] contrast. The Bengali inventory contains a clerical error, too. According the source (Ray et al. 1966: 4), the low back vowel is "rounded and low," which is [ɒ]. If so, the contrast between [æ a] is in backness, not in height. The Gujarati inventory in P-base is [i u e ə o æ a ɑ]. However, in a more recent study, Cardona and Suthar (2003: 723) list the vowels as [i u e ə o ɛ ɔ a], where [a] is the only low vowel. The Maithili inventory contains another clerical error. According to the source (Yadav 1996: 14), [æ a] are the only unrounded low vowels, which contrast in backness, not height. We conclude that there is no compelling evidence for three-way contrasts in backness at any vowel height, or for a height contrast among low vowels.

3.2.2 *Pharyngeal vowels*

As discussed above, a pharyngeal vowel has a narrowed pharynx, achieved by retracting the tongue root, or [−ATR]. Therefore, we ask whether pharyngeal vowels can contrast with [−ATR] vowels. A search through P-base yields two languages that contain pharyngeal vowels, Tsakhur and !Xoo (spelled as !Xu in UPSID); in neither language is there a contrast between pharyngeal vowels and [−ATR] vowels. We conclude therefore that there is no basis for keeping both "pharyngeal" and [−ATR].

3.2.3 Creaky vs. glottalized vowels

Creaky vowels (such as [a̰]) involve some kind of glottal closure, so do glottalized vowels (such as [aˀ]). Therefore, we ask whether the two types contrast in any language. A search in P-base yields no such contrast and we conclude that there is no need to keep both terms.

3.2.4 Voiceless vowels

Voiceless vowels occur in whispered speech and in certain contexts of normal speech, such as in an unstressed syllable with an aspirated onset (e.g. *potato* [pʰə̥teto] and *suppose* [sə̥pʰoz]). What remains to be seen is whether voiceless vowels can contrast with other vowels in regular contexts. A search in P-base yields one inventory that contains voiceless vowels, Turkana. However, according to the source (Dimmendaal 1983: 31–3), voiceless vowels only occur in special conditions, summarized in (34).

(34) Conditions on voiceless vowels in Turkana (Dimmendaal 1983: 31–3)
   a. Limited to final position.
   b. Limited to polysyllabic words.
   c. Limited to high vowels (low voiceless vowels are deleted or "subtracted").
   d. Preceding vowel is longer than usual.

An alternative analysis, offered by Polley and Jeffrey (1977), is to consider a devoicing process conditioned by stress: when a final vowel is stressed, it is voiced, otherwise it is devoiced (or deleted). The stress analysis can account for two additional facts: (a) a monosyllable is not devoiced and (b) when a "voiceless" (i.e. unstressed) vowel is not final (and hence voiced), it is still shorter than its preceding vowel. It is worth noting that in a later study, Dimmendaal (1993: 131) accepted the marginal status of voiceless vowels and excluded them from the phoneme inventory of Turkana.

3.2.5 Extra-short and extra-long vowels

Next we consider extra-short and extra-long vowels. If they are real, we need up to a four-way contrast in length, such as [ă a aː aːː]. A search for extra-short vowels in P-base yields two languages, Ostyak (also called "Khanty") and Vietnamese. Both are found in UPSID and have been discussed earlier. A search for extra-long vowels in P-base also yields two languages, Mixe (Guichicovi) and Wichita. The P-base inventory of Mixe (Guichicovi) has a clerical error: it is the phonetic (allophonic) inventory, not the phoneme inventory. According to the source (Wichmann 1995: 56), "the inventory of vowels can be reduced to /i i: e e: ï ï: ɨ a a: u u: o o:/," which shows a two-way contrast in length only.

Wichita is found in both UPSID and P-base. However, the inventory is given differently in the two databases. This is shown in (35).

(35) Vowel inventory of Wichita, as given in UPSID and P-base

Database	Inventory	Source
UPSID	[i eː ɛ ɛː u oː ɒ ɒː]	Garvin (1950), Rood (1975)
P-base	[i iː iːː e eː eːː a aː aːː]	Rood (1976)

It is worth noting that the same author, David Rood, is consulted in both studies (the phonology part of Rood 1976 is the same as Rood 1975). Besides, according to Rood (1976), there has been "only one dialect" of Wichita since the 1930s, and Garvin (1950) is a "thorough description" of the phonetics of Wichita. Nevertheless, while P-base contains extra-long vowels, UPSID does not. The difference is that, while Rood (1975; 1976) prefers to use length to represent most vowel contrasts, Garvin (1950) considers some contrasts to lie in quality (height and backness or rounding). In fact, Rood is aware of the quality difference among the vowels. For example, he makes the comments in (36).

(36) Some observations on Wichita vowels (Rood 1976: 229–30)
/iː iːː/ are lower than /i/.
/eː eːː/ are lower than /e/.
[iːː] is rare; [aːː] is even rarer; [e] is rare but [eːː] is common.
[u o] do occur.

If the rare forms [iːː aːː e] are not full phonemes, then Wichita only has two degrees of length. If we attribute some contrasts to height, we do not need three degrees of length either. Either way, the case is weak for a three-way contrast in length. It is worth noting that Rood (1975) is advocating the "singular" nature of Wichita, which includes (i) the absence of back vowel and (ii) a three-way length contrast. Neither property seems conclusive. For example, following his notes, /e/ can be realized as [æ] and /a/ ranges from central to back, which means that [æ ɑ] (or [æ ɒ]) do occur. In addition, if the "glide" /w/ is [u], there are two high vowels [i u]. Rood's analysis is not impossible, but equally good or better alternatives are available. We conclude, therefore, that there is no compelling evidence for extra-short or extra-long vowels.

3.2.6 Other vowels

Finally, we consider a set of remaining vowels that involve a diacritic for a less common feature. They are listed in Table 3.14.

A "fronted" vowel has a slightly more forward tongue position than its regular counterpart. They are found in three languages, shown in Table 3.15.

In Quechua, the fronted [a] does not contrast with a regular [a]. In Swedish, the fronted [ʉ] does not contrast with a regular [ʉ] or with [ʏ] (if it is a front vowel, as McClean 1947 describes it); we shall return to Swedish in Chapter 6. In

TABLE 3.14 Vowels that involve an extra diacritic for a less common feature and the number of inventories in which such a vowel is found

	Fronted	Advanced	Retroflex	Apical	Muffled	Lowered	Raised
Example	[ə̟]	[ɔ̘]	[ɚ]	[i̺]	[emf]	[i̞]	[a̝]
Inventories	3	1	1	1	1	9	14

TABLE 3.15 Languages that contain "fronted" vowels, as given in P-base

Cuzco Quechua	[ɪ ʊ a̟]
Swedish	[iː yː ʉ̟ː i y ɵ ʊ eː øː ə oː ɛ œ ɔ a ɑː]
Thompson	[i̟ i u e ə̟ ə o a]

Thompson, [i̟ i] could be interpreted as [i ɪ] and [ə̟ ə] could be interpreted as [ə ɤ]. Therefore, a feature for "fronted" is not necessary in any of the languages.

The feature "advanced" is found in Af Tunni Somali only. It refers to the tongue root position, which is similar to [+ATR], or what are sometimes called "tense" vowels. Therefore, we ask whether there is a contrast between "advanced" and [+ATR] in Af Tunni Somali, and the answer is no.

Retroflex vowels use both the tongue tip and the tongue body. Since the tongue body and the tongue tip are independent articulators, they are free to occur together. In addition, it is easier to add a tongue tip gesture to a back vowel, because there is more room for the tongue tip movement. In P-base, only one language has retroflex vowels, which are [ɨ˞ ɚ] in Mising, neither being front.

There is one "apical" vowel in P-base, found in Bisu. We have seen Bisu in UPSID, where the apical vowel is a syllabic consonant.

There are two "muffled" vowels in P-base, found in Grebo, which has seven nasal vowels [ĩ ũ ẽ õ ɛ̃ ɔ̃ ã] and nine oral vowels [i u emf omf e o ɛ ɔ a], where [mf] means "muffled". According to the source (Innes 1966: 14), the term "muffled" is borrowed from Sapir (1931) in his study of Gweabo, "a language closely related to Grebo." Gweabo has eleven vowels, five being "bright" and six being "muffled." There is also a harmony process called "bright to bright, muffled to muffled." A possible analysis is to interpret "muffled" vowels as [+ATR] and "bright" vowels as [–ATR] (Singler 1983: 7). There are two reasons for the analysis. First, "muffled" vowels have a slightly higher tongue position than "bright" vowels. Second, there are more "muffled" vowels in Gweabo than "bright" vowels (six vs five) (Sapir 1931: 31).

Nine inventories in P-base have "lowered" vowels. (We have excluded Saami, which contains a clerical error.) The data are summarized in Table 3.16.

Table 3.16 Inventories that contain "lowered" vowels, as given in P-base

German (Michigan)	[ai iː ĩː uː au ɪ ʊ eː ẽː oː õː ɛ̞ ə ɔ̞ æː aː ãː ɑː a ɑ]
Haitian Creole	[ĩ ũ i̞ u̞ ẽ õ e o ɔ̃ ɛ ɔ a]
Kedang	[i i̞ u u̞ ɛ ɛ̞ o o̞ æ æ̞ a a̞]
Kalenjin	[i u ɪ ʊ e o ɛ ɔ a a̞]
Maya (Itzaj)	[i u ɨ̞ e o a iː uː eː oː aː]
Maya (Chontal)	[i u ɨ̞ e o a]
Tangale	[i u e o a i̞ u̞ e̞ o̞]
Welsh (2 varieties)	[i ɪ ɨ ɨ̞ u ʊ e ɛ o ɒ a ɑ ə]

Table 3.17 Inventories that contain "raised" vowels, as given in P-base

Akan (5 varieties)	[i u e̝ o̝ e o ɛ ɔ a̝ a ĩ ũ ẽ̝ õ̝ ã]
Chrau	[iː uː ɪ ʊ e̝ː o̝ː e o ɛː ɔː a̝ aː]
Khmer	[i ɯ u e̝ ə̝ e ə o ɛ ɔ ɑ ɒ]
Kimatuumbi	[i̝ u̝ i u ɛ ɔ a]
Lama	[i ɨ u ɪ ʌ̝ ʊ e o ɛ ɔ a]
Mam	[ɪ ʊ ɛ ɔ a iː uː eː oː a̝ː]
Shilluk	[i̝ u̝ i u e̝ ɔ̝ e ɔ a̝ a i̝ː u̝ː iː uː e̝ː ɔ̝ː eː ɔː a̝ː aː]
Tzutojil (3 varieties)	[ɪ ʊ ɛ ɔ a̝ iː uː eː oː aː ie uo]

A "lowered" vowel has a slightly lower tongue position than its regular counterpart. We ask whether the difference is contrastive. In Michigan German, there is no contrast between [ɛ̞ ɔ̞] and [ɛ ɔ]. In Haitian, there is no contrast between [i̞ u̞] and [i u]. In Kedang, there is a contrast between "lowered" vowels and their regular counterparts. However, the source, Samely (1991: 11), considers the contrast to be one of voice quality, one being "modal" and the other "breathy." Therefore, Kadeng does not call for a new feature. In Kalenjin, [a a̞] could be [ʌ a]. In Maya (both varieties), [ɨ] could be [ə]. In Tangale, [i̞ u̞ e̞ o̞] could be [ɪ ʊ ɛ ɔ]. Finally, in Welsh (standard and northern), a natural interpretation of the contrast between [ɨ ɨ̞] is one of tense–lax (or ATR) as well, which could be transcribed as [ɨ ᵻ], as it is in the source (Thorne 1993: 4), or [ɯ ɯ] if we count "central" vowels as back.

Finally, we consider "raised" vowels, found in fourteen inventories. The data are summarized in Table 3.17. For the five varieties of Akan, we have chosen the longest inventory as the example. The same is true for the three varieties of Tzutojil.

In many cases, a "raised" vowel does not contrast with its regular counterpart. For example, in Akan, [ẽ̝ õ̝] do not contrast with [ẽ õ]. In other cases, a "raised" vowel does not contrast with a regular vowel in its neighborhood. For example, in Akan, [ɪ ʊ] are close neighbors of [e̝ o̝] but not used. Based on these considerations, we can interpret the inventories without assuming "raised" vowels. This is shown in Table 3.18.

Table 3.18 Analysis of the inventories in Table 3.17, without using "raised" vowels

Akan	[i u ɪ ʊ e o ɛ ɔ ʌ a ĩ ũ ẽ õ ã]
Chrau	[iː uː ɪ ʊ eː oː ɛ ɔ æː ɒː ʌ aː]
Khmer	[i ɯ u ɪ ɯ e ə o ɛ ɔ ɑ ɒ]
Kimatuumbi	[i u ɪ ʊ ɛ ɔ a]
Lama	[i ɨ u ɪ ʌ ʊ e o ɛ ɔ a]
Mam	[ɪ ʊ ɛ ɔ ʌ iː uː eː oː aː]
Shilluk	[i u ɪ ʊ e o ɛ ɔ ʌ a iː uː ɪː ʊː eː oː ɛː ɔː ʌː aː]
Tzutojil	[ɪ ʊ ɛ ɔ ʌ iː uː eː oː aː]

For some inventories, there may be more than one alternative analysis. For example, in Chrau, if we consider length to be distinctive for some vowel pairs, for example treating [iː ɪ] as [iː i], we can simplify the inventory to seven vowels [i u e o æ a ɔ], and account for length separately. Similarly, Haiman (2011: 1–2) considers length to be distinctive in Khmer, and lists the vowel phonemes as [i e ɛ æː ɯ ɤ ɔ a u o ɑ], where every vowel can be long or short except [æː], which alternates with a diphthong [ae:] and is always long.

This completes our discussion of vowels with special diacritics. As far as contrast is concerned, we have found none of them to be necessary.

3.3 Summary

Our examination of UPSID and P-base has found no three-way contrast in backness. In addition, there is only a two-way contrast in height (or in tenseness or ATR) among high or mid vowels, and no contrast in height (or in tenseness or ATR) among low vowels. Therefore, there are at most twenty basic vowels, shown in Table 3.19, where only nineteen are found.

Table 3.19 The inventory of basic vowels in UPSID and P-base

		Front		Back	
High	Regular (tense)	i	y	ɯ	u
	Lower (lax)	ɪ	ʏ	ɯ	ʊ
Mid	Regular (tense)	e	ø	ɤ	o
	Lower (lax)	ɛ	œ	ʌ	ɔ
Low		æ		ɑ	ɒ

Non-basic vowels can be created by adding additional features to a basic vowel, such as nasality, creakiness, or murmur. In addition, a retroflex feature can be added to a back vowel. Moreover, an additional timing slot can create diphthongs and long vowels. The expansion ends there, though. For example, we found no compelling evidence for extra-short vowels, extra-long vowels, voiceless vowels, or fricative vowels, not because they cannot be pronounced, but because no language is found to use them contrastively.

4

Vowel height

We have seen in Chapter 3 that, if we consider tenseness or ATR to be a separate feature, then no language requires more than three degrees of contrast in tongue height for vowels. However, tongue height remains the only feature that has three degrees of contrast. It is possible that tongue height in fact comprises two binary features [high] and [low], as proposed by Chomsky and Halle (1968). However, it is worth asking, in the first place, whether all three degrees of height are needed in any language, a question that has not been raised before. If no language requires three degrees of contrast in tongue height, then four binary features—[high], [back], [round], and [ATR] (or [tense])—are sufficient to distinguish all basic vowels. Specifically, if the four features can combine freely, there are just sixteen basic vowels, shown in Table 4.1.

Among high vowels, the higher one of a pair is [+ATR] (e.g. [i] of [i ɪ] and [u] of [u ʊ]), evidenced by X-ray studies (Ladefoged 1964; Stewart 1967; Halle and Stevens 1969; Perkell 1969; 1971; Wood 1982; Lindau 1979; Kenstowicz 2009). In addition, we follow Halle and Stevens (1969: 212) and consider "low" vowels to be [–ATR]. One may note that while [ATR] corresponds to [tense] for high vowels, it does not

TABLE 4.1 A two-height system of sixteen basic vowels, using four binary features [high], [back], [round], and [ATR]. Since there are more IPA symbols than needed, the symbol choice for some cells is somewhat arbitrary

		–back		+back	
		–round	+round	–round	+round
+high	+ATR	i	y	ɯ	u
	–ATR	ɪ	ʏ	ʌ	ʊ
–high	+ATR	e	ø	ɜ	o
	–ATR	ɛ	œ	ɑ	ɔ

A Theory of Phonological Features. First edition. San Duanmu.

for low vowels. For example, if [+tense] is defined by occurrence in a stressed open vowel, such as [ɑ] in *spa* and [ɔ] in *law* in American English, then [ɑ] and [ɔ] should be [+tense], yet they are shown as [–ATR] in Table 4.1. This means that the feature [tense] in English requires some reconsideration, to which we shall return.

A two-height system may seem rather radical. Therefore, I shall start with motivations for questioning the traditional system. Then I examine vowel contrast in UPSID and P-base and see whether any language requires three degrees of height. Next, I examine evidence from sound classes and see if a two-height system is sufficient. Finally, I consider the case of incremental "step raising" (Parkinson 1996) and discuss how it can be analyzed in a two-height system.

4.1 Motivations for a two-height system

Let us consider two motivations for questioning the standard assumption that (at least) three-degrees of vowel height are needed: the existence of two-height languages and the most common vowel systems.

In some languages, while the vowels seem to show three (or more) degrees of phonetic height, two degrees of phonological height are sufficient to distinguish them. A well-known example is Turkish. The phonetic properties of its eight vowels were seen in Fig. 1.2 and are repeated here in Fig. 4.1, and the phonological properties of the vowels are shown in (1).

(1) Phonological properties of Turkish vowels

	–back		+back	
	–round	+round	–round	+round
+high	i	y	ɯ	u
–high	e	œ	a	o

Phonetically, Turkish vowels seem to show three degrees of height, where [a] is low (and central or front). However, phonologically, two degrees of height are sufficient (Lewis 1967), where [a] can be grouped with [e œ o] in height (and with [ɯ u o] in

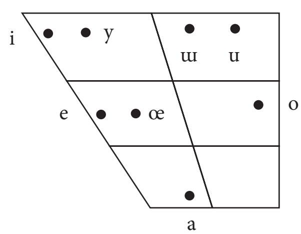

FIG. 4.1 Phonetic properties of Turkish vowels (Zimmer and Orgun 1992: 44).
Reproduced with permission from Cambridge University Press.

Table 4.2 Top three most common vowel inventories in P-base, which account for 23 per cent of the 628 inventories

Inventory	Count	Sample language
[i u e o a]	87	Spanish
[i u a]	38	Yupik
[i u e ɛ o ɔ a]	24	Italian

backness). This can be seen in the alternation between two forms of the plural suffix [ler] and [lar], where [ler] is used after a front vowel and [lar] after a back vowel. If [e a] have the same height, the alternation involves a change of backness only, conditioned by the preceding vowel. If [e a] differ in height as well, the analysis is more complicated, and unnecessarily so.

Next we consider the most common vowel systems. The top three in P-base are shown in Table 4.2. Similar patterns can be seen in UPSID.

It can be seen that in all the cases, there is a lack of height contrast between a mid vowel and a low vowel. In particular, [a] is the most common low vowel, which in most cases can be analyzed as an unrounded non-high back vowel, similar to the case in Turkish. For example, a two-height analysis of Italian is shown in (2).

(2) Two-height analysis of Italian vowels

		–back	+back	
		–round	–round	+round
+high		i		u
–high	+ATR	e		o
	–ATR	ɛ	a	ɔ

It remains to be seen whether all inventories can be represented in a two-height system, as far as contrast is concerned. Two problematic cases are shown in (3) and (4).

(3) Basic vowels in Pacoh, as given in UPSID (converted to IPA symbols)

High	i	ɨ	u
Mid tense	e	ɘ	o
Mid lax	ɛ	ɜ	ɔ
Low	æ	a	ɒ

(4) Basic vowels in Woisika, as given in Stokhof (1979: 59) (original symbols)

	Front		Central		Back	
	Lax	Tense	Lax	Tense	Lax	Tense
High	i	î			u	û
Mid	e	ê			o	ô
Low	ae		a	â	ao	
	Unrounded				Rounded	

Pacoh has fewer than sixteen basic vowels, which in principle can fit into a two-height system. However, it is not obvious how this could be achieved. For example, we need to analyze either [e] or [ɛ] as [+high, –ATR], but it is not obvious which one should be chosen, or how it could be justified. The basic vowels in Woisika present a different challenge: There are five unrounded front vowels and five rounded back vowels, whereas a two-height system only has four corresponding positions for each set.

In what follows, I shall examine all such inventories in UPSID and P-base in order to find out which ones present genuine problems for the two-height system. I shall consider evidence from both contrast and sound classes.

4.2 Basic vowels in UPSID

Out of the 451 inventories in UPSID, just thirty-four have sixteen or more vowels. In addition, all of them contain non-basic vowels. For illustration, let us consider Irish, which has twenty-four vowels, shown in (5).

(5) Vowels in Irish, as given in UPSID
[ɪ e a o ʊ ĩ ẽ ã õ ʊ̃ iː eː ɑː oː uː ĩː ẽː ɑ̃ː õː ũː əi əu iə uə]

The system can be analyzed with five basic vowels [i e ɑ o u], which can be short (transcribed as [ɪ e a o ʊ]) or long. Alternatively, the system can be analyzed with eight basic vowels, high [i u ɪ ʊ] and non-high [e a ɑ o], where [e o] can be long or short, [i ɑ u] are always long, and [ɪ a ʊ] always short.

If we exclude non-basic vowels, no language has more than fifteen vowels. Let us consider all inventories that have twelve or more basic vowels. There are seven such inventories, summarized in Table 4.3.

In Po-Ai, seven of its fifteen basic vowels are "extra-short." As discussed in Chapter 3, not all of them are independent phonemes. Instead, Po-Ai has just nine vowel phonemes, which do not present a problem for a two-height system.

The Khanty vowel inventory is [i y ɯ u e ö o æ a ə ö̆ ă ŏ], of which [ə ö̆ ă ŏ] are "reduced" vowels (Gulya 1966: 23). If we treat them as unstressed versions of other

Table 4.3 Inventories with twelve or more basic vowels in UPSID

Language	All vowels	Basic vowels
Po-Ai	16	15
Khanty	13	13
Bete	13	13
Dan	20	12
Pacoh	15	12
Sebei	12	12
Woisika	12	12

vowels, Khanty only has nine basic vowels, which can easily fit into a two-height system. Alternatively, if we treat reduced vowels as [−ATR], the inventory can be analyzed in (6), where [ə] is probably an unstressed vowel and should be excluded from the inventory.

(6) Two-height analysis of basic vowels in Khanty (pairs separated by comma differ in [ATR])

	−back		+back	
	−round	+round	−round	+round
+high	i	y	ɯ, ə	u
−high	e, æ	ö, ŏ̈	a, ă	o, ŏ

Next we consider Bete. According to Werle and Gbalehi (1976), the UPSID source, Bete has thirteen vowels, shown in (7).

(7) Vowels in Bete (Werle and Gbalehi 1976: 61)

		−back	+back	
			−round	+round
+high	+ATR	i	ɨ	u
	−ATR	ɩ	ɜ	ʊ
−high	+ATR	e	ə	o
−low	−ATR	ɛ	ʌ	ɔ
+low		a		

[a] is the only vowel that does not fit into a two-height system. It is odd in two other ways. First, [a] is unspecified for [ATR] (or for [back] or [round]), a feature that divides other vowels into two harmony groups. Second, there is a lack of contrast between [ə] and [a]. Specifically, Werle and Gbalehi (1976: 61–7) give two word pairs for each pair of contrastive vowels, but the word pairs for [ə]-[a] are not minimal pairs, since they differ in tone. This is shown in (8), where tones are indicated by a diacritic over the vowel.

(8) "Minimal pairs" for the [ə]-[a] contrast in Bete (Werle and Gbalehi 1976: 65)

[kpə́]	*chaise* ('chair')	[gwə́]	*fesses* ('buttocks')
[kpa̍]	*banco* ('bank')	[gwā]	*attacher* ('attach')

In the first pair, [kpə́] has a high tone but [kpa̍] has a mid-high. In the second pair, [gwə́] has a high tone but [gwā] has a mid-low. If the contrast between [ə] and [a] cannot been established, Bete is not a compelling case against a two-height system.

Next we consider Dan. According to the source (Bearth and Zemp 1967: 19–21), Dan has four degrees of height and three degrees of backness, where front and central vowels are unrounded and back vowels rounded. There are twelve basic vowels (plus eight nasal vowels), shown in (9).

(9) Basic vowels in Dan (Bearth and Zemp 1967: 19–21)

	Front	Central	Back
Close	i	ɨ	u
Mid close	e	ɘ	o
Mid open	ɛ	ə	ɔ
Open	æ	a	ɒ

This inventory can be analyzed in (10), where pairs separated by a comma differ in [ATR] and where we have kept the original IPA symbols.

(10) Two-height analysis of basic vowels in Dan (pairs separated by a comma differ in [ATR])

	–back	+back	
		–round	+round
+high	i, e	ɨ, ɘ	u, o
–high	ɛ, æ	ə, a	ɔ, ɒ

The authors comment that "mid close vowels tend to be higher" (Bearth and Zemp 1967: 20), which supports our analysis that treats them as high lax vowels.

Next we consider Pacoh, cited earlier, which also has twelve basic vowels (each can be long or short). The UPSID source, Watson (1964), divides them into three degrees of height. However, in a more recent study, Watson (1996) only assumes two degrees of height, each of which in turn is divided into two by the feature [ATR]. This is shown in (11). In the analysis, Dan differs from Pacoh only in the choice of IPA symbols for some of the vowels.

(11) Basic vowels in Pacoh (Watson 1996)

		–back	+back	
			–round	+round
+high	+ATR	i	ɯ	u
	–ATR	i̙	ɯ̙	u̙
–high	+ATR	e	ɤ	o
	–ATR	e̙	ɤ̙	o̙

Watson (1996: 203) offers an interesting explanation of why he started out with three degrees of height and ended with just two: it was "partly due to training in tongue-height articulation which only gradually gave way to an understanding of tongue-root articulation and phonation types."

Next we consider Sebei, whose vowels are [i e a ɔ o u ï ë ä ɔ̈ ö ü], of which [ï ë ä ɔ̈ ö ü] are "short" versions of the others. If they contrast in length, then Sebei only has six basic vowels. On the other hand, the short vowels do not all seem to be full phonemes. According to the source (Montgomery 1970: 51), "the distribution of short vowels is limited to word-medial occurrences," "/ä/, /ö/ and /ɔ̈/ occur rarely," and "contrasts within the set of short vowels have a low functional yield." It is possible then that the short vowels are unstressed versions of other vowels, although Montgomery does not discuss stress. It is safe to exclude at least [ä ɔ̈ ö], since they are rare and do not contrast with other short vowels. The remaining vowels are analyzed in (12).

(12) Two-height analysis of basic vowels in Sebei (pairs separated by a comma differ in [ATR])

	–back	+back	
		–round	+round
+high	i, ï		u, ü
–high	e, ë	a	o, ɔ

Finally, let us consider Woisika, whose vowels are repeated in (13), based on Stokhof (1979: 59).

(13) Basic vowels in Woisika (Stokhof 1979: 59) (original features and symbols)

	Front		Central		Back	
	Lax	Tense	Lax	Tense	Lax	Tense
High	i	î			u	û
Mid	e	ê			o	ô
Low	ae		a	â	ao	
	Unrounded				Rounded	

Two comments can be made of the inventory. First, as Stokhof (1979) points out, tense vowels are long (and have a lower tone). This can be seen in the spectrograms (Stokhof 1979: 135–9), where a tense vowel is twice as long as a lax one. If length is contrastive (although Stokhof 1979: 61 does not think so), we can exclude the five "tense" (long) vowels and the rest can fit into a two-height system. Second, we note that [ae ao] are unspecified for tenseness. If we can account for [ae ao] separately, then the remaining vowels can fit into a two-height system. This is shown in (14), where pairs in a cell differ in [tense] (or probably [ATR]).

(14) Two-height analysis of basic vowels in Woisika, excluding [ae ao]

	–back	+back	
	–round	–round	+round
+high	î, i		û, u
–high	ê, e	â, a	ô, o

There is some evidence for excluding [ae ao]. First, they are quite rare. Consider the frequency data provided by Stokhof (1979: 114), shown in Table 4.4, based on a text of 11,302 phoneme tokens.

We see that [ae ao] have the lowest frequencies (along with [û]). Therefore, it is possible they are not independent phonemes. There is another reason for assuming a smaller vowel inventory. In UPSID, Woisika is classified as a Trans-New Guinea language. There are altogether twenty-six languages in the group, and their average number of basic vowels is six. Woisika has not only the largest inventory of basic vowels in the group, but three more than the runner-up. If we exclude [ae ao], Woisika still has the largest inventory of basic vowels in this group.

In summary, while some vowel inventories in UPSID seem to require more than two degrees of height, such as Bete and Woisika, none is found to be conclusive. We conclude that, as far as clear evidence is concerned, a two-height system is sufficient to distinguish all contrasts among basic vowels.

TABLE 4.4 Vowel frequencies in Woisika, based on a text of 11,302 phoneme tokens (Stokhof 1979: 114), after excluding consonants

Vowel	Count	%
/â/	1,639	31.5
/e/	1,074	20.6
/i/	801	15.4
/a/	588	11.3
/o/	455	8.7
/u/	262	5.0
/ê/	164	3.1
/ô/	158	3.0
/î/	50	1.0
/û/	9	0.2
/ae/	9	0.2
/ao/	1	0.0
All	5,210	100.0

4.3 Basic vowels in P-base

Of the 628 inventories in P-base, thirty-three have seventeen or more vowels. If we look at basic vowels only, then just one inventory has more than sixteen vowels. As in the previous section, we examine all inventories that contain twelve or more basic vowels. There are seventeen such inventories, summarized in Table 4.5.

In Chapter 3 we discussed Turkana, whose voiceless vowels are not full phonemes; in a more recent work the original author, Dimmendaal (1993: 131), proposes just nine vowels, which can easily fit into a two-height system.

The German inventory, based on Fox (1990), has five unrounded front vowels [i ɪ e ɛ æ], exceeding the four slots in a two-height system. In addition, there are three back unrounded non-high vowels [ə a ɑ], whereas a two-height system has just two such slots. However, according to Kohler (1999: 87), [ɛ æ] contrast in length, analyzed as [ɛ ɛː], so do [a ɑ], analyzed as [a aː]. In Kohler's analysis, German has fourteen basic vowels, which can be analyzed in (15), where pairs separated by a comma differ in [ATR].

(15) Two-height analysis of basic vowels in German

	–back		+back	
	–round	+round	–round	+round
+high	i, ɪ	y, ʏ		u, ʊ
–high	e, ɛ	ø, œ	ə, a	o, ɔ

Table 4.5 Seventeen inventories in P-base that have twelve or more basic vowels. The three varieties of Dutch have the same basic vowels, as do the three varieties of Welsh.

Language	Basic vowels
Turkana	18 [i u i̥ u̥ ɪ ʊ ɪ̥ ʊ̥ e o e̥ o̥ ɛ ɔ ɛ̥ ɔ̥ a ḁ]
German	16 [i y u ɪ ʏ ʊ e ø o ɛ œ ə ɔ æ a ɑ]
Danish	13 [i y u ɪ ʊ ø ʌ o ɛ œ ɔ æ a]
Dutch (3 varieties)	13 [i y u ɪ ʏ e ø o ɛ œ ɔ a ɑ]
Karimojong	13 [i y u ɪ ɵ ʊ e o ɛ œ ɔ a ɑ]
Ostyak (Khanty)	13 [i y ɯ u e ø o æ a ə̆ ø̆ ă ŏ]
Saami	13 [i y ɨ ɜ u ɨ̞ e ʌ ɵ o ə ɐ ɑ]
Welsh (3 varieties)	13 [i ɨ u ɪ ɨ̞ ʊ e o ɛ ə ɒ a ɑ]
American English	12 [i ɪ ə ʊ u e ɛ ʌ ɔ o æ a]
Louisiana Creole French	12 [i y u ɪ e ø o ɛ œ ɔ æ a]
Khmer	12 [i ɯ u e̝ ə̝ e ə o ɛ ɔ ɑ ɒ]
Tibetan	12 [i y u ɪ ʊ e ø o ɛ ɔ ə a]
Tsakhur	12 [i iˤ u uˤ e eˤ ə əˤ o oˤ a aˤ]

Danish has thirteen basic vowels, most of which can be long or short. The analysis is shown in (16), where pairs separated by a comma differ in [ATR].

(16) Two-height analysis of basic vowels in Danish

	−back		+back	
	−round	+round	−round	+round
+high	i, ɪ	y		u, ʊ
−high	ɛ, æ	ø, œ	ʌ, a	o, ɔ

The basic vowels in Dutch (three varieties) present no problem either. The analysis is shown in (17), where again pairs separated by a comma differ in [ATR].

(17) Two-height analysis of basic vowels in Dutch

	−back		+back	
	−round	+round	−round	+round
+high	i, ɪ	y, ʏ		u
−high	e, ɛ	ø, œ	a, ɑ	o, ɔ

Next we consider Karimojong. If we interpret the central vowel [ɵ] as front, the analysis is shown in (18).

(18) Two-height analysis of basic vowels in Karimojong

	–back		+back	
	–round	+round	–round	+round
+high	i, ɪ	y		u, ʊ
–high	e, ɛ	ɵ, œ	a, ɑ	o, ɔ

Ostyak is also called Khanty. It is included in UPSID (from the same source) and has been analyzed earlier, so it is not repeated here.

The P-base inventory of Saami contains a clerical error, as discussed earlier. Otherwise, Saami has just six vowel phonemes, in agreement with UPSID, which also contains Saami from the same source.

Next we consider Welsh. According to the source (Thorne 1993: 4), of the thirteen vowels, [i ɨ u e o ɑ] "are phonologically long" and [ɪ ɨ̞ ʊ ɛ ə ɒ a] "phonologically short." This means that we only need to assume up to seven basic vowels, which present no problem for a two-height system.

Next we consider American English. We note that [ə] does not contrast with [ʌ], because [ə] is always unstressed and [ʌ] stressed. We also note that some vowels are long (and can be a diphthong, such as [e o]). Therefore, American English has at most eleven basic vowels, which can be analyzed in (19), where a comma indicates a contrast in [ATR].

(19) Two-height analysis of basic vowels in American English

	–back	+back	
	–round	–round	+round
+high	iː, ɪ		uː, ʊ
–high	eː, ɛ æː	ʌ aː	oː, ɔ

The analysis does not present the only possible solutions. For example, [ʌ a] could differ in length (as shown), or in [high] (not shown). It can be seen that the feature [tense] is not well reflected in (19). We shall return to this point below.

Next we consider vowels in Louisiana Creole French, which is also called "Louisiana Creole" or "Louisiana French." The point of interest is that, in its inventory of basic vowels [i y u ɪ e ø o ɛ œ ɔ æ a] (from Klingler 1992: 40–42), there are five unrounded front vowels [i ɪ e ɛ æ], whereas in a two-height system there are just four relevant slots. However, several issues are worth noting. First, [ɪ] is marginal and not used in the author's later work (Klingler 2003: 148). Second, [e ɛ] "alternate freely" and rarely contrast with each other. Third, [æ] is mostly used in English words or as an allophone of [ɛ] before [r]. Finally, in other studies, fewer basic vowels are proposed. For example, in Valdman et al. (1998: 5), of which Flingler is a co-author, only ten basic vowels are proposed, where [i e ɛ] are the only unrounded front vowels. If we follow Valdman et al. (1998), the analysis is as in (20).

(20) Two-height analysis of basic vowels in Louisiana Creole French

	–back		+back	
	–round	+round	–round	+round
+high	i	y		u
–high	e, ɛ	ø, œ	a	o, ɔ

The analyses of the remaining three languages are given in (21)–(23), where we interpret "raised" mid vowels in Khmer to be [+high, –ATR] and pharyngealized vowels in Tsakhur as [–ATR].

(21) Two-height analysis of basic vowels in Khmer

	–back		+back	
	–round	+round	–round	+round
+high	i, ę̝		ɯ, ə̝	u, o
–high	e, ɛ		ə, ɑ	ɔ, ɒ

(22) Two-height analysis of basic vowels in Tibetan

	–back		+back	
	–round	+round	–round	+round
+high	i, ɪ	y		u, ʊ
–high	e, ɛ	ø	ə, a	o, ɔ

(23) Two-height analysis of basic vowels in Tsakhur

	–back		+back	
	–round	+round	–round	+round
+high	i, iˤ		ə, əˤ	u, uˤ
–high	e, eˤ		a, aˤ	o, oˤ

This concludes our analysis of basic vowels in P-base. As in UPSID, a two-height feature system is sufficient to distinguish all contrasts among basic vowels.

4.4 Evidence from sound classes

The preceding discussion is based on contrast, where a two-height system is shown to be sufficient. Let us now consider whether the system is able to account for sound classes. Specifically, we consider two common features that are absent in the two-height system, [tense] in English and [+low] in P-base.

4.4.1 [tense] in English

In English, [tense] is often defined by occurrence in stressed open syllables (e.g. Ladefoged and Johnson 2011: 99). Thus, [i u e o ɑ ɔ ɝ ai au oi] in American English are [+tense], since they occur in such syllables, and [ɪ ʊ ɛ æ ʌ] are [−tense]. However, as Ladefoged and Johnson (2011: 310) point out, [tense] has "no specific phonetic correlates."

Halle and Stevens (1969) propose that [tense] corresponds to [ATR]. This works for high vowels, where [i u] have advanced tongue root and [ɪ ʊ] do not (shown in Fig. 4.2, based on cineradiography). There is some evidence, too, that [e o] have

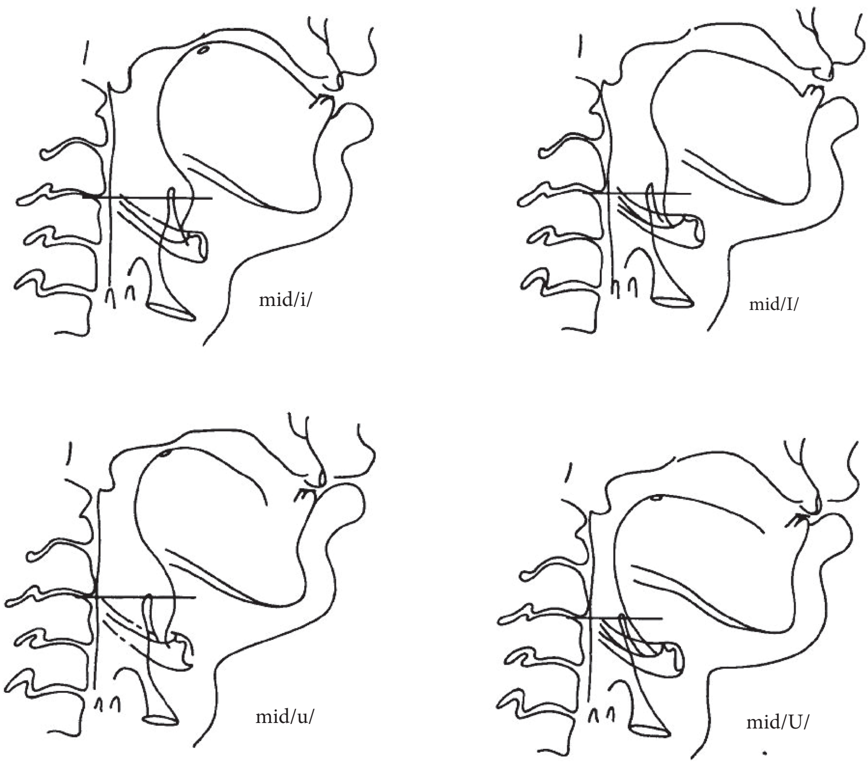

Fig. 4.2 Difference in ATR (advanced tongue root) between [i ɪ] (top panels) and between [u ʊ] (bottom panels), from Halle and Stevens (1969: 210).

advanced tongue root and [ɛ ʌ] do not (Perkell 1971; Wood 1982). On the other hand, as Halle and Stevens note, low vowels do not have advanced tongue root, yet [ɑ ɔ] are thought to be [+tense].

Perkell (1971) interprets [+low] as [+constricted pharynx], and [+tense] as either [+ATR] or [+constricted pharynx], where [tense] is still not a homogeneous class. Ladefoged and Johnson (2011: 98) suggest that the difference between tense and non-tense vowels is partly "due to developments in the history of the English language that are still represented in the spelling." This means that vowel distributions in English do not always imply genuine sound classes, and [tense] may not be a genuine phonological feature either.

Yet another possibility is that [tense] corresponds to vowel length, where tense vowels are long and non-tense vowels short. Consider a two-height analysis of monophthongs in American English, shown in (24). Vowel length is based on Ladefoged and Johnson (2011: 90), except [æ], which is phonetically long too (Peterson and Lehiste 1960). [e: o:] are also included, although they are sometimes realized as diphthongs [ei ou].

(24) Two-height analysis of monophthongs in American English

+ATR	i:			u:
−ATR	ɪ			ʊ
+ATR	e:		ɝ:	o:
−ATR	ɛ æ:		ʌ ɑ:	ɔ:

In this analysis, all tense vowels are long (so are diphthongs, which are also tense). The vowel [æ] seems exceptional: it is phonetically long but does not occur in open syllables. However, [æ] does occur in an open syllable in a few marginal words. In addition, in New York–Philadelphia English, [æ] undergoes "tensing" and becomes a diphthong (Benua 1995), which is expected if [æ] is long in the first place. In fact, once length is made use of, not all speakers need to treat [ɛ ʌ], or even [ɪ ʊ], as [−ATR]. Such cross-speaker variability has been noted by Ladefoged et al. (1972) and Lindau (1979).

In summary, the traditional feature [tense] is problematic for any theory. Its analysis in a two-height system is at least as good as in other approaches. Therefore, there is no need to add additional features, such as [tense] or [low].

4.4.2 [+low] in P-base

P-base contains a large collection of sound classes, each of which is supposed to share one or more common features. Let us consider whether the feature [low] is needed for any of them. To do so, I propose the procedure in (25).

(25) Procedure for collecting and checking sound classes that involve [+low]
a. Collect all sound classes that contain one or more of [æ a ɑ ɔ ɒ].
b. Exclude classes that contain a non-low vowel.
c. Exclude classes that contain a consonant.

First, a sound class without a low vowel is not defined by [+low]. Since [ɔ] is sometimes used as a mid vowel and sometimes a low vowel, I group it with low vowels, in order not to miss potentially relevant cases. Next, we exclude classes that contain a non-low vowel, since they cannot be defined by [+low]. Finally, it is reasonable to assume that [+low] is not specified for consonants; if a sound class contains a consonant, the class is not defined by [+low] either.

There are 9,039 sound classes in P-base, including repetitions of the same class (i.e. those whose members are identical). Our procedure yields just thirty-five classes of interest. Eighteen of the classes consist of one basic low vowel, sometimes plus variations of it. For example, consider the classes in (26).

(26) Sound classes that consist of one basic "low" vowel

Class	*Features (at most)*
[a]	[+back, –round, –high, –nasal]
[a ã]	[+back, –round, –high]
[ɔ ɔ̃]	[+back, +round, –high]

If a class contains a single basic vowel, or variations of it, there is no need for [+low]. The reason is that, as seen in the preceding section, every vowel inventory in P-base can be interpreted in a two height system. Therefore, any basic vowel can be unambiguously referred to without using [+low].

The remaining seventeen classes come from fifteen languages (including two from Boko/Busa and two from Sri Lanka Creole Portuguese). They are summarized in Table 4.6.

Thirteen of the classes consist of [ɔ a] or [ɔ ɔ̃ a ã]. In six languages (Arbore, Bemba, Muna, Russian, North Slavey, and South Slavey), [ɔ a] can be defined as [+back, –high]. In Boko/Busa (two identical classes), [ɔ ɔ̃ a ã] can also be defined as [+back, –high]. Next, we consider Kpelle, the analysis of which is shown in (27).

(27) Analysis of vowels in Kpelle (pairs in a cell differ in [ATR])

	–back		+back	
	–round	+round	–round	+round
+high	i			u
–high	e, ɛ		a	o, ɔ

Table 4.6 Sound classes made of two or more "low" vowels in P-base, along with their language names and vowel inventories

Language	Class	Inventory
Arbore	[ɔ a]	[i u ɛ ɔ a]
Bemba	[ɔ a]	[i u ɛ ɔ a]
Ejagham	[ɔ a]	[i y u ɛ ə ɔ a]
Kana (Khana)	[ɔ a]	[ĩ ũ i u e o ɛ̃ ɔ̃ ɛ ɔ ã a]
Kpelle	[ɔ a]	[i u e o ɛ ɔ a]
Muna	[ɔ a]	[i u ɛ ɔ a]
Russian	[ɔ a]	[i ɨ u ɛ ɔ a]
North Slavey	[ɔ a]	[i u e ɛ ɔ a]
South Slavey	[ɔ a]	[i u ɛ ɔ a]
So (Soo)	[ɔ a]	[i u ɪ ʊ e o ɛ ɔ a]
Temne	[ɔ a]	[i ɨ u e ʌ o ɛ a ɔ]
Boko/Busa	[ɔ ɔ̃ a ã]	[i ĩ u ũ e o ɛ ɛ̃ ɔ ɔ̃ a ã]
Boko/Busa	[ɔ ɔ̃ a ã]	[i ĩ u ũ e o ɛ ɛ̃ ɔ ɔ̃ a ã]
Chamorro	[æ a]	[i u e o æ a]
Mixe, Midland	[æ a]	[i ɨ u e o æ a]
SLC Portuguese	[æ ɔ æː aː]	[i u e ə o æ ɔ iː uː eː oː æː aː ɔː]
SLC Portuguese	[æ ɔ]	[i u e ə o æ ɔ iː uː eː oː æː aː ɔː]

It is reasonable to assume that [o] is [+ATR]. If [a] is [−ATR], we can define [ɔ a] as [+back, −high, −ATR]. A similar analysis can be offered for So (Soo). In Temne and Ejagham, if [ʌ] or [ə] is [+ATR], we can again define [ɔ a] as [+back, −high, −ATR]. Finally, we can define [ɔ a] in Kana (Khana) as [+back, −high, −ATR, −nasal].

There are four classes left, from three languages (Chamorro, Midland Mixe, and Sri Lanka Creole Portuguese). The analysis of Chamorro is shown in (28).

(28) Analysis of vowels in Chamorro (pairs in a cell differ in [ATR])

	−back		+back	
	−round	+round	−round	+round
+high	i			u
−high	e, æ		a	o

We can define [æ a] as [−high, −ATR]. Interestingly, Chamorro also has a class [e o], which can be defined as [−high, +ATR].

Next we consider Midland Mixe. Since its vowel inventory is similar to that of Chamorro, we can define [æ a] as [−high, −ATR], too.

Finally, consider Sri Lanka Creole (SLC) Portuguese. There are seven short vowels and seven long vowels, which are analyzed in (29).

(29) Analysis of vowels in SLC Portuguese (pairs in a cell differ in [ATR])

	Short			Long		
	–back	+back		–back	+back	
	–round	–round	+round	–round	–round	+round
+high	i		u	iː		uː
–high	e, æ	ə	o, ɔ	eː, æː	aː	oː, ɔː

Again, if [e o ə] are [+ATR] and [æ ɔ a] [–ATR], then we can define [æ ɔ] as [–high, –ATR] and short.

The final class, [æ ɔ æː aː] in Sri Lanka Creole Portuguese, presents a problem. If we define it as [–ATR], the class should be [æ ɔ æː aː ɔː], where [ɔː] is also a member, not in the original class. However, the problem has little to do with a two-height system. Even if we assume a three-height system, the problem is still there. This is shown in (30).

(30) Analysis of vowels in SLC Portuguese in a three-height system

	Short			Long		
	–back	+back		–back	+back	
	–round	–round	+round	–round	–round	+round
High	i		u	iː		uː
Mid	e	ə	o	eː		oː
Low	æ		ɔ	æː	aː	ɔː

In a three-height analysis, [æ ɔ] can be defined as [+low] and short, but there is no way to define [æ ɔ æː aː]: if we define it as [+low] (long or short), the class is [æ ɔ æː aː ɔː], where [ɔː], not in the original class, is also included.

Classes like [æ ɔ æː aː] are not uncommon in P-base; Mielke (2008) calls them "unnatural classes" because they cannot be defined by any feature set. The explanations of "unnatural classes" are not relevant for the present discussion. It suffices to know that adding the feature [low] does not solve the problem; therefore there is no reason for doing so.

4.5 Step raising

"Step raising" refers to cases where vowels of different heights are raised by one step each. Parkinson (1996) discusses a number of such cases and argues that vowel height should be represented incrementally, rather than with binary features. As an example, let us consider Lena Spanish (Hualde 1989: 785–90) and compare three analyses. For convenience, I call them (i) the incremental-height analysis (Parkinson 1996), (ii) the three-height analysis, which uses two binary features [high] and [low], and (iii) the two-height analysis, which uses one binary feature [high].

Lena Spanish has five vowels [i u e o a]. In some context, [u] can raise the preceding [a] to [e] and [e o] to [i u]. The data are shown in (31), from Hualde (1989: 785–90).

(31) Step raising in Lena Spanish (Hualde 1989)

	Fem. sg.	*Masc. sg.*	*Gloss*
[a] → [e]	gáta	gétu	'cat'
	sánta	séntu	'saint'
[e] → [i]	néna	nínu	'child'
	bwéna	bwínu	'good'
[o] → [u]	kósa	kúsu	'cripple'
	bóna	búnu	'good'

In the incremental-height analysis (Parkinson 1996: 19–20), height is represented by how many tokens of the feature [closed] a vowel has. The representation of the vowels in Lena Spanish is shown in (32). An additional feature can distinguish [i] from [u] and [e] from [o], which we omit.

(32) Incremental-height analysis of vowels in Lena Spanish (Parkinson 1996)

[i]	[u]	[e]	[o]	[a]
|	|	|	|	
[closed]	[closed]	[closed]	[closed]	
|	|			
[closed]	[closed]			

The more tokens of [closed] a vowel has, the higher the vowel. The lowest vowel [a] is unspecified for [closed]. The prose form of the step-raising rule is given in (33). Since the target vowel gains a token of [closed], it is raised by one step in height.

(33) Incremental-height analysis of step raising in Lena Spanish
If [u] follows a vowel X, where X has fewer tokens of [closed] than [u], spread the lowest token of [closed] from [u] to X.

Next we consider the three-height analysis, shown in (34) and (35), which uses two traditional binary features [high] and [low] (Chomsky and Halle 1968).

(34) Three-height analysis of vowels in Lena Spanish

Vowels	*Features*
[i u]	[+high, –low]
[e o]	[–high, –low]
[a]	[–high, +low]

(35) Three-height analysis of step raising in Lena Spanish
[+low] → [–low] / __ [+high, –low]
[–high, –low] → [+high] / __ [+high]

Three comments are in order. First, in the three-height analysis, step raising involves two assimilation rules, instead of one; we shall return to this point. Second, to prevent the rules from applying to [a] twice, another condition is needed, which we omit here. Third, other analyses are possible, such as using [high] and [ATR] (instead of [high] and [low]), or specifying [a] as [+low] without [–high]. The alternatives are not considered here, because they have little consequence for our discussion.

Finally, let us consider the two-height analysis, shown in (36) and (37). We follow the Successive Division Algorithm of Dresher (2009) and assume that a language can choose any set of features to distinguish its vowels. In the case of Lena Spanish, the features are [ATR], [high], and [round], in that order. Other ways of specifying the vowels are of little consequence and are not discussed here.

(36) Two-height analysis of vowels in Lena Spanish

	[i]	[u]	[e]	[o]	[a]
[ATR]	+	+	+	+	–
[high]	+	+	–	–	
[round]	–	+	–	+	

(37) Two-height analysis of step raising in Lena Spanish
[–ATR] → [+ATR] / __ [+high, +ATR]
[–high] → [+high] / __ [+high]

Let us now compare the three analyses. It is clear that the two-height analysis is better than the three-height one. First, they both use two rules for step raising, so the three-height analysis has no advantage. Second, we have seen that a two-height system is sufficient to distinguish all vowels. Therefore, the three-height system is unnecessarily complicated, and over-predicts possible vowel contrasts.

Next we compare the two-height analysis with the incremental-height analysis, where the former is again superior. First we note that incremental heights alone cannot account for true [ATR] harmony, where [ɪ] can change to [i] and [ɛ] to [e] but [e] cannot change to [ɪ]. In addition, [ATR] harmony is not always triggered by a higher vowel; for example, the change of [ɪ] to [i] can be triggered by [e]. Therefore, the incremental-height theory still needs the feature [ATR], as acknowledged by Parkinson (1996: 73). If so, the theory has considerable redundancy. In particular, apart from [ATR], Parkinson (1996: 7–8) proposes at least four degrees of height, yet we have seen that two degrees are sufficient. Why then does Parkinson favor multiple heights in the analysis of Lena Spanish?

The only argument given in Parkinson (1996: 20) is that step raising ought to be analyzed with a single rule, or a "unified description," probably because it is traditionally referred to by a single term "metaphony" and is triggered by a single vowel [u]. The argument has several problems. First, it is not clear, even in the incremental-height analysis, whether the change from [a] to [e] is the same as that from [e] to [i] (or [o] to [u]), because there is a structural difference: In the former, [closed] is spread directly to the vowel, whereas in the latter, it is spread to an existing token of [closed]. Second, in the two-rule analysis, each change involves one feature from the same source [u], which could be viewed as a "unifying" property, if it is at all important. Third, if the phonetic transcription is accurate, [a] is [–ATR] and [i u e o] [+ATR], and the change from [a] to [e] does involve a change in [ATR], whereas that from [e o] to [i u] does not. Therefore, the two-height analysis is phonetically accurate, whereas the incremental-height analysis is not. Finally, in the incremental-height analysis, height is an odd case among all features: it is the only one that allows the same feature to stack up on top of each other, whereas no other feature does. Such a dramatic change in feature theory requires compelling evidence. If alternative analyses are unavailable or unnatural, the justification would be strong. However, as we have seen, alternative analyses are readily available, fully natural (in terms of assimilation), and possibly more accurate phonetically. Therefore, there is no advantage in adopting incremental heights.

4.6 Summary

I have shown that all contrasts among basic vowels in UPSID and P-base can be represented in a two-height system, which uses four binary features [back], [high], [round], and [ATR], giving a total of sixteen basic vowels. In addition, an examination of all sound classes in P-base shows that there is no need for [low]. Moreover, I have discussed different analyses of step raising and shown that the two-height system offers the simplest solution. We conclude that a two-height vowel system is a viable and simpler alternative to traditional theories of vowel features.

5

Consonant contrasts

In this chapter I examine consonant contrasts in UPSID and P-base. I discuss consonants in UPSID first, followed by those in P-base, where similar cases are discussed only briefly.

5.1 Consonants in UPSID

UPSID contains 650 consonant types (distinct transcriptions), with a total of 10,133 tokens (two UPSID transcriptions are excluded: "G<" is not found and "h2" is the same as "hh"). They can be divided into a set of "basic" consonants and other sets that involve an additional gesture or airstream mechanism. The categories are shown in Table 5.1, where the category "X-ized" includes various secondary articulations, such as labialized, palatalized, and velarized. If a consonant falls under two (or more) categories, it is arbitrarily included in just one. For example, a pre-nasalized affricate is grouped with "Pre-nasalized" and not with "Affricate." Because every category is accounted for, it does not matter whether a pre-nasalized affricate is grouped with "Pre-nasalized" or "Affricate." The same is true for other consonants that fall into two or more categories.

Some categories require little elaboration. In particular, I assume that clicks, ejectives, and implosives involve different airstream mechanisms (Ladefoged and Johnson 2011), but their other gestures are similar to those of basic consonants. In addition, a long consonant can be represented as a regular one linked to two timing slots (Chapter 1). Moreover, "Laryngealized" (also called "Glottalized" in Maddieson 1984: 99) is a glottal gesture independent from the oral closure; therefore, this feature is not controversial either. Finally, a consonant can use two (or more) gestures, as long as they belong to separate articulators. For example, breathy and aspirated consonants involve a glottal gesture, independent from oral gestures. Similarly, labial-velar consonants involve two independent gestures, so are consonants in the "X-ized" category, such as [t^w], where a second gesture (the secondary articulation) [^{w}] is added to that of a basic consonant [t]. In what follows, I shall focus on basic consonants. We return to pre-nasalized consonants, nasally released consonants (also called post-nasalized consonants), and affricates in Chapter 7.

There are 134 basic consonants. Of these, 128 can be placed in an extended IPA table, shown in Table 5.2. The places of articulation, as named in UPSID, are Bilabial (BL), Labio-Dental (L-D), Dental (Den), unspecified for Dental or Alveolar (D/A), Alveolar

A Theory of Phonological Features. First edition. San Duanmu.

Table 5.1 Consonants in UPSID. Categories with * are discussed in Chapter 7

Category	Type	Token
Basic	134	6,979
Affricate*	47	703
Click/Ejective/Implosive	135	628
X-ized	169	509
Labial-velar	6	429
Breathy/Aspirated	35	428
Pre-nasalized*	51	251
Laryngealized	40	150
Long	27	50
Nasally released*	6	6
All	650	10,133

(Alv), Retroflex (Retr), Palato-Alveolar (P-A), Palatal (Pal), Velar (Vel), Uvular (Uvu), Pharyngeal (Pha), and Glottal (Gl). The manners of articulation are Plosive, Nasal, Trill, Tap, Flap, Sibilant Fricative (Sib Fric), Fricative (Fric), Lateral Fricative (L-Fric), Lateral Flap (L-Flap), Approximant (Appr), Lateral Approximant (L-App), and r-sound. When a manner has two lines, the first is voiced and the second voiceless.

Since UPSID lists more sounds than are found in a regular IPA table, I have created some new symbols. For example, I use [*] to indicate sounds that are ambiguous between dental and alveolar, and [$^{\mathrm{NS}}$] to indicate some non-sibilant fricatives. Other raised letters indicate either the place of articulation, such as [j$^{\mathrm{PA}}$] (a palatal-alveolar [j]) or the manner of articulation, such as [β$^{\mathrm{App}}$] (an approximant [β]). Given their locations in the table, the meanings of the new symbols should be transparent.

There are six remaining "basic" consonants, which are listed in Table 5.3. The feature descriptions are those used in UPSID.

For each consonant, we ask whether it can contrast with a more common sound. The result is shown in Table 5.4.

In Guranti, there is no contrast between a fricative trill and a regular trill. In Kiowa, there is no contrast between a "pre-stopped lateral" and a lateral affricate (to be discussed later with affricates). In Ekari, there is no contrast between a regular [g] and a "laterally released" [g]. In Axluxlay, there is no contrast between a "voiceless velar-alveolar lateral fricative" and [l̥] with velar aspiration; nor is there a contrast between [k] with "alveolar lateral fricative release" and [k$^{\mathrm{l}}$]. Both [l̥$^{\mathrm{x}}$] and [k$^{\mathrm{l}}$] are complex sounds that involve two separate articulators, to be discussed later. Finally, in Tacana, there is no contrast between a fricative flap and an aspirated flap. In fact, an examination of the sources shows that this sound should be a palatalized flap, or [ɾ$^{\mathrm{j}}$] (Van Wynen and Van Wynen 1962: 17; Key 1968: 28).

We now return to basic consonants in Table 5.2. We focus on issues that seem to be exceptional, which are given in (1).

TABLE 5.2 128 basic consonants in UPSID; raised symbols, such as [*] and [^{NS}] are explained in the text

	BL	L-D	Den	D/A	Alv	Retr	P-A	Pal	Vel	Uvu	Pha	Gl
Plosive	b	b̪	d̪	d*	d	ɖ	d^{PA}	ɟ	g	ɢ	ɢ^{Ph}	ʔ̬
	p		t̪	t*	t	ʈ	t^{PA}	c	k	q		ʔ
Nasal	m	m̪	n̪	n*	n	ɳ	n^{PA}	ɲ	ŋ	N		
	m̥		n̪̥	n̥*	n̥	ɳ̥	n̥^{PA}	ɲ̥	ŋ̥			
Trill			r̪	r*	r	r^{Ret}		r^{Pal}		R		
				r̥*	r̥							
Tap			ɾ̪^{Tp}		ɾ^{Tp}							
Flap		v	ɾ̪	ɾ*	ɾ	ɽ	ɾ^{PA}					
Sib Fric			z̪	z*	z	ʐ	ʒ	ʑ				
			s̪	s*	s	ʂ	ʃ	ʑ̥				
Fric	β	v̪	ð	z^{NS*}	z^{NS}	ʐ^{NS}	ʒ^{NS}	ʝ	ɣ	ʁ	ʕ	ɦ
	ɸ	f̪	θ			ʂ^{NS}	ʃ^{NS}	ç	x	χ	ħ	h
L-Fric			ɮ̪	ɮ*	ɮ	ɭ^{Fr}						
			ɬ̪	ɬ*	ɬ				ʟ̥^{Fr}			
L-Flap				ɺ^{Fl*}	ɺ^{Fl}	ɭ^{Fl}						
Appr	β^{App}	ʋ		ɹ*	ɹ	ɻ	j^{PA}	j	ɰ	ʁ^{App}		
	ɸ^{App}			ɹ*				j̥				
L-App			l̪	l*	l	ɭ	l^{PA}	ʎ	L			
			l̪̥	l̥*	l̥	ɭ̥						
r-sound			r̪r̪	rr*	rr							
				r̥r̥*								

TABLE 5.3 Remaining six "basic" consonants that do not fit into Table 5.2

UPSID	Token	Description
"rF	1	Voiced dental/alveolar fricative trill
dlD	1	Voiced pre-stopped dental lateral approximant
gL	1	Laterally released voiced velar plosive
hxlF	1	Voiceless velar-alveolar lateral fricative
klF	1	Voiceless velar plosive with alveolar lateral fricative release
r[F	1	Voiced alveolar fricative flap

TABLE 5.4 Analysis of the consonants in Table 5.3

UPSID	Language	Analysis
"rF	Gurani	No contrast with a regular trill [r]
dlD	Kiowa	No contrast with lateral affricate [dl]
gL	Ekari	No contrast with [g]
hxlF	Axluxlay	No contrast with [l̥ˣ]
klF	Axluxlay	No contrast with [kˡ]
r[F	Tacana	No contrast with [ɾʰ] or [ɾʲ]

(1) Issues to be examined among basic consonants:
Voiced glottal stop [ʔ̬] and pharyngeal stop [ɢᴾʰ]
Contrast among dental, dental/alveolar, and alveolar places
Contrast between sibilant and non-sibilant fricatives
Contrast between taps and flaps
Contrast between lateral fricatives and lateral approximants
Contrast between r-sounds and other approximants

5.1.1 Voiced glottal stop and pharyngeal stop

The voiced glottal stop is reported in Karen, which is a clerical error. An examination of the source (Jones 1961: 5) shows that this sound is a regular voiceless glottal stop.

The voiced pharyngeal stop is reported in three languages: Avar, Bats, and Iraqw. According to Ladefoged (1971: 41), it is impossible to make such a sound, nor is it found in the standard IPA table. An examination of the inventories shows that none of the languages has a contrast between a voiced pharyngeal stop and a voiced uvular stop, the latter being a known sound. A close look at Iraqw is instructive. In Maddieson (1984: 315), Iraqw has a pharyngeal stop and a uvular stop, both voiceless. However, in UPSID, updated by the same author (Maddieson and Precoda 1990), the pharyngeal stop is voiced, while the uvular stop is voiceless; if so, both stops can be interpreted as uvular, contrasting in voice rather than in place. Moreover, in P-base, Iraqw has no pharyngeal stop, based on a newer source (Nordbustad 1988). In summary, there is no compelling evidence for assuming pharyngeal stops.

5.1.2 Dental, dental/alveolar, and alveolar places

Next we consider the places of dental, dental/alveolar, and alveolar. We ask whether all three places are needed. To answer it, we search for all possible contrasts among them. The result is shown in (2).

(2) Contrasts among dental (D), dental/alveolar (DA), and alveolar (A):

Contrast	Example	Result
D-DA-A triplets	[θ s* s]	Not found
D-DA pairs	[θ s*]	Not found
DA-A pairs	[s* s]	Not found
D-A pairs	[θ s]	Found

There is no contrast between D and DA or between DA and A. Therefore, we can reduce "Dental or alveolar" by merging it with "Alveolar" (there is little consequence if we merge DA with D instead).

5.1.3 *Sibilant vs. non-sibilant fricatives*

Nine pairs differ in this feature. A search through UPSID shows four pairs that are contrastive. In all the cases, however, an alternative analysis is available. This is shown in (3), without the feature sibilant.

(3) Contrastive pairs between sibilant and non-sibilant fricatives

Pair	Found in	Analysis
[z̪ ð]	Albanian	[z ð]
[s̪ θ]	Albanian, Spanish	[s θ]
[ʒ ʒ^{NS}]	Kabardian	[ʒ ʑ]
[ʃ ʃ^{NS}]	Kabardian	[ʃ ɕ]

In Albanian, the source considers the contrast to be in place, not in sibilance. In particular, [ð θ] are "inter-dental" and [z̪ s̪] are "post-dental" (Newmark 1957: 17–19). In addition, there is no contrast between [z̪ s̪] and [z s]. Therefore, we can consider the contrast to be interdental for [θ ð] and non-interdental for [s z]. A similar analysis can be given for Spanish.

In Kabardian, the source considers the contrast to be in place, too. In particular, Kuipers (1960: 18) calls [ŝ ẑ] "palatal-alveolar" and [š ž] "alveolar-palatal," and we can interpret them as [ɕ ʑ] and [ʃ ʒ] respectively. In Chapter 6, I shall discuss palatals in more detail and suggest that the tongue tip is raised in [ʒ ʃ] but not in [ʑ ɕ].

5.1.4 *Taps and flaps*

Next we consider whether a tap and a flap can contrast with each other. UPSID lists two such pairs: [ɾ̪ ɾ̪^{Tp}] and [ɾ ɾ^{Tp}]. However, a search through UPSID shows that neither pair is contrastive in any language. Therefore, we can merge tap and flap into a single category.

We also consider the contrast between a lateral flap and non-lateral flap. A search yields one pair, which is found in three languages, shown in (4).

(4) Contrast between a non-lateral flap [ɾ] and lateral flap [l$^{\text{Fl}}$]

Language	Comment	Analysis
Paez	No [r] or [l]	[r l]
Yucuna	No [r] or [l]	[r l]
Kewa	No [r] or [l]	[r l]

None of the languages has a contrast between a flap [r] and a regular [r], or between a flap [l] and a regular [l]. Therefore, the contrastive pair could be represented as [r l].

5.1.5 Lateral fricatives vs. lateral approximants

While both "lateral" and "fricative" are known features, their combination is rather unusual, because laterals are typically sonorants while fricatives are not. Therefore, we want to find out whether lateral fricatives can be confirmed. There are seven relevant pairs in Table 5.2. A search through UPSID found three pairs, shown in Table 5.5.

Diegueno has four laterals [l ɬ l̪ ɮ̪], divided into two places (alveolar and dental). The alveolar pair can be distinguished by voice, where [l] is voiced and [ɬ] voiceless. The dental pair can be distinguished by aspiration, where [l̪] is unaspirated and [ɮ̪] aspirated (breathy). In neither case do we need to assume a lateral fricative.

The other nine languages contain three basic laterals each (we ignore non-basic laterals here, such as long laterals, labialized laterals, and lateral affricates). In three of the languages, the place is unspecified between dental and alveolar. In the other six languages, the place is indicated as alveolar. Let us use [l ɬ ɮ] to represent the basic laterals in all of them. Two analyses are available without assuming lateral fricatives. One makes use of [voice] and [aspiration], shown in (5), where [+voice, +aspirated] is also called "breathy."

Table 5.5 Contrastive pairs between a lateral fricative and a lateral approximant in UPSID; [*] indicates a place that could be dental or alveolar

UPSID	IPA	Found	Languages
[lDF]-[lD]	[ɮ̪]-[l̪]	1	Diegueno
[hlDF]-[hlD]	[ɬ̪]-[l̪̥]	0	
["lF]-["l]	[ɮ*]-[l*]	3	Archi, Ik, Socotri
[lF]-[l]	[ɮ]-[l]	6	Kanakuru, Lame, Margi, Ngizim, Tera, Zulu
["hlF]-["hl]	[ɬ*]-[l̥*]	0	
[hlF]-[hl]	[ɬ]-[l̥]	0	
[l.F]-[l.]	[ɭ$^{\text{Fr}}$]-[ɭ]	0	

(5) Analysis of basic laterals in terms of [voice] and [aspirated]

Lateral	[l]	[l̥]	[ɬ]	[ɮ]
[voice]	+	−	−	+
[aspirated]	−	−	+	+

In the second analysis, a "fricative lateral" has an additional palatal gesture (by raising the tongue body towards the palate). This is shown in (6).

(6) Analysis of basic laterals in terms of secondary articulation

Lateral	[l]	[l̥]	[ɬ]	[ɮ]
[voice]	+	−	−	+
Palatal	−	−	+	+

There are two arguments for (6). First, while the basic laterals in Archi are transcribed with the same place in UPSID (either dental or alveolar), Chumakina et al. (2008: 2) split them into two places: [l] is dental and [ɬ ɮ] are palatal-alveolar. Similarly, Ladefoged and Maddieson (1996: 206) transcribe the laterals as [l L L̥], where [L L̥] are "pre-velar," which is a similar region to palatal. The second argument comes from Ngizim (Schuh 1972: 6), where all alveolar consonants are palatalized before [i], but not [l ɬ ɮ] (or [r r̃]). This restriction is natural if [ɬ ɮ] are already palatal; in addition, to palatalize [l] would change it to [ɮ], which explains why there is no contrast between a palatalized [l] and [ɮ].

In summary, there is no compelling evidence for the category of lateral fricatives. Instead, known features, such as [voice], [aspirated], and a secondary palatal articulation are sufficient to distinguish all basic laterals in UPSID.

5.1.6 "r-sound" and "approximant [r]"

Three such pairs are found in Table 5.2. A search through UPSID yields the result in Table 5.6. Since no contrast is found, we can omit the category "r-sound."

5.1.7 Summary

An exhaustive examination of UPSID shows that many listed contrasts are not substantiated. Instead, Table 5.2 can be reduced to Table 5.7.

Table 5.6 Contrastive pairs between an "approximant [r]" and an "r-sound" in Table 5.2; a search through UPSID found no such pair

UPSID	IPA	Found
["rA]-["rr]	[ɹ*]-[rr*]	0
["hrA]-["hrr]	[ɹ̥*]-[r̥r̥*]	0
[rA]-[rr]	[ɹ]-[rr]	0

Table 5.7 Eighty-nine basic consonants in UPSID, after removing non-contrastive sounds from Table 5.2.

	BL	L-D	ID	D/A	Retr	P-A	Pal	Vel	Uvu	Pha	Gl
Plosive	p b	b̪	t̪ d̪	t d	ʈ ɖ	t^{PA} d^{PA}	c ɟ	k g	q ɢ		ʔ
Nasal	m̥ m	ɱ	n̪̊ n̪	n̥ n	ɳ̊ ɳ	ɲ̊ ɲ	ŋ̟̊ ŋ̟	ŋ̊ ŋ	N		
Trill			r̪	r̥ r	r^{Ret}		r^{Pal}		R		
Flap/Tap		v	ɾ̪	ɾ	ɽ	ɾ^{PA}					
Fricative	ɸ β	f̪ v̪	θ ð	s z	ʂ ʐ	ʃ ʒ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Appr	ɸ^{App} β^{App}	ʋ	ɹ̪	ɹ̥ ɹ	ɻ	j^{PA}	j̊ j	ɰ	ʁ^{App}		
L-App			l̪̊ l̪	l̥ l	ɭ̊ ɭ	l^{PA}	ʎ	ʟ̥ ʟ			

As we shall see shortly, this table is sufficient for representing all contrasts among basic consonants in P-base, too.

5.2 Consonants in P-base

P-base contains 786 consonant types (distinct transcriptions), with a total of 14,645 tokens. We again divide them into different categories, shown in Table 5.8. As with UPSID, if a consonant has properties of two (or more) categories, it is arbitrarily included in just one. For example, [k͡p^w] is grouped with "double" and not with "secondary."

Table 5.8 Consonants in P-base. Categories with * are discussed in this chapter and those with ** in Chapter 7

Categories	Example	Type	Token
Basic*	[m]	63	9,843
Secondary	[d^j d^w]	246	1,732
Click, Ejective, Implosive	[! t' ɓ]	172	762
Affricate**	[bv ts]	45	752
Pre- and post-nasalized**	[^{n}d t^n]	63	322
Placeless nasal	[N?]	1	8
Others*	[ȁ ʔk]	72	266
Double	[k͡p]	47	789
Long	[b:]	45	85
Breathy	[m̤ z̤]	6	14
Voiceless*	[m̥ d̥]	15	33
Nasalized	[w̃ r̃]	7	33
Syllabic	[m̩]	4	4
Total		786	14,645

The category "Basic" refers to those that can fit into a traditional IPA table. The category "Secondary" refers to those that have a raised diacritic, traditionally called secondary articulation; the category also includes aspirated consonants, such as [tʰ]. The category "Double" refers to doubly articulated consonants, i.e. those with two equal oral closures, including [w]. "Breathy" refers to aspiration in voiced consonants. "Voiceless" refers to a diacritic (circle below) that is added to what are otherwise voiced consonants. "Nasalized" refers to nasalization of voiced continuants (except [ɦ̃], which seems to be an error for [ɦ]; see Elugbe 1989: 91). "Placeless nasal" refers to nasals whose place is determined by the following consonant in a NC cluster. The terms "Affricate," "Click," "Ejective," "Implosive," "Pre-nasalized," "Post-nasalized," "Long," and "Syllabic" are familiar in the literature. The category "Others" refer to those with other diacritics.

We shall leave affricates, pre- and post-nasalized stops, clicks, ejectives, and implosives for Chapter 7. In addition, we assume that a consonant can use two (or more) independent articulators at the same time. Therefore, no elaboration is needed for secondary articulations (which normally involve an articulator different from the primary one), doubly articulated sounds, or nasalized continuants. A placeless nasal will acquire its place from the following consonant and become one of the nasals in the "basic" category, so no separate discussion is needed. Breathy consonants involve voiced aspiration, which is not controversial, but we should check whether there is a contrast between two alternative diacritic symbols, such as [m̤ mʱ], and no such contrast is found. Finally, a long consonant can be represented by a regular one linked to two timing slots (Chapter 1) and a syllabic consonant can be represented by its location in a syllable; therefore, these categories require no elaboration either. In what follows, we focus on the categories "Basic," "Voiceless," and "Others."

5.2.1 Basic consonants

The sixty-three basic consonants in P-base are shown in Table 5.9. The places of articulation are Bilabial (BL), Labio-dental (L-D); Interdental (ID); Dental (Den); Alveolar (Alv); Retroflex (Ret), Alveolar-Palatal (A-P), Pre-palatal (PPal), Palatal (Pal), Velar (Vel), Uvular (Uvu), Pharyngeal (Pha), and Glottal (Gl). The manners of articulation are Stop, Nasal, Trill, Flap, Fricative (Fric), Approximant (App), and Lateral approximant (L-App). When a cell has two sounds, the first is voiceless and the second voiced.

Two properties of this table require some discussion: (i) the four-way contrast among interdental, dental, alveolar, and retroflex places, and (ii) the three-way contrast among alveolar-palatal, pre-palatal, and palatal places.

There are no interdental stops, but there are seven triplets among dental, alveolar, and retroflex places. They are shown in (7), along with the search results.

TABLE 5.9 Sixty-three basic consonants in P-base. Voiceless sonorants, such as [m̥], are grouped under "Voiceless" (see Table 5.8) and not included here. Aspirated consonants, such as [t^{h}], are grouped under "Secondary" (see Table 5.8) and are not included here either

	BL	L-D	ID	Den	Alv	Ret	A-P	PPal	Pal	Vel	Uvu	Pha	Gl
Stop	p b			t̪ d̪	t d	ʈ ɖ			c ɟ	k g	q ɢ		ʔ
Nasal	m		n̪		n	ɳ			ɲ	ŋ	ɴ		
Trill					r						ʀ		
Flap		ⱱ		ɾ̪	ɾ	ɽ							
Fric	ɸ β	f v	θ ð	s̪ z̪	s z	ʂ ʐ	ʃ ʒ	ɕ ʑ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
App				r̪	ɹ	ɻ			j	ɰ			
L-App				l̪	l	ɭ			ʎ				

(7) Seven triplets in Table 5.9 for dental, alveolar, and retroflex places

Triplet	[t̪ t ʈ]	[d̪ d ɖ]	[s̪ s ʂ]	[z̪ z ʐ]	[ɾ̪ ɾ ɽ]	[l̪ l ɭ]	[r̪ ɹ ɻ]
Languages	12	2	0	0	0	8	0

The data show that a three-way contrast among dental, alveolar, and retroflex places is possible for stops (and laterals), but not for fricatives (or rhotics). Let us take a close look at pairs of contrastive places for fricatives, shown in (8).

(8) Pairs of contrastive places for fricatives

Place pair	[θ s̪]	[ð z̪]	[s̪ s]	[z̪ z]	[s̪ ʂ]	[z̪ ʐ]	[s ʂ]	[z ʐ]
Languages	2	2	0	0	4	0	13	3

The point of interest here is that fricatives do not contrast between dental and alveolar places, although stops do. On the other hand, fricatives do contrast between interdental and dental places, but stops do not. To account for these properties, I propose a hypothesis in (9) which I call the Coronal Hypothesis, since the tongue tip (coronal) is involved in all four places (interdental, dental, alveolar, and retroflex).

(9) The Coronal Hypothesis
The tongue tip (coronal) can make use of up to three contrastive places (e.g. dental, alveolar, and retroflex for stops, or interdental, alveolar, and retroflex for fricatives).

The Coronal Hypothesis is consistent not just with data in P-base but with those in UPSID, which we discussed above.

Let us now consider contrasts among alveolar-palatal, pre-palatal, and palatal places. There are two such triplets in P-base, shown in (10).

(10) Two triplets in Table 5.9 for alveolar-palatal, pre-palatal, and palatal places

Triplet	[ʃ ɕ ç]	[ʒ ʑ ʝ]
Languages	0	0

None of the languages in P-base seems to make use of all three kinds of palatals. On the other hand, there are languages that make use of two palatal places. This is shown in (11).

(11) Pairs of two kinds of palatal places in Table 5.9

Place pair	[ʃ ɕ]	[ʒ ʑ]	[ʃ ç]	[ʒ ʝ]	[ɕ ç]	[ʑ ʝ]
Languages	4	3	3	4	1	0

While contrasts between different kinds of palatals seem infrequent, there are in fact a lot more palatal sounds, many of which are represented with a secondary articulation, not included in the table of basic consonants. I shall return to this issue in Chapter 7.

This completes our discussion of basic consonants in P-base. In Table 5.10 we revise Table 5.9 slightly.

5.2.2 Consonants with a "voiceless" diacritic

Fifteen consonants have a voiceless diacritic [r̥ m̥ n̥ ʎ̥ j̊ l̥ ɲ̊ ɥ̊ ŋ̊ b̥ d̥ ɟ̊ ɡ̊ ɣ̊ ʒ̊]. It is uncontroversial that sonorant consonants (nasals, liquids, and glides) can be voiced or voiceless. Therefore, we focus on obstruent consonants (stops and fricatives).

TABLE 5.10 Sixty-one basic consonants in P-base, reduced from Table 5.9, after (i) removing the Dental column by shifting some sounds to Interdental and some to Dental/Alveolar (D/A), and (ii) deleting [s̪ z̪], which do not contrast with [s z]. This table is nearly identical to that obtained from UPSID, except that the UPSID table comprises eighty-nine consonants

	BL	L-D	ID	D/A	Ret	A-P	PPal	Pal	Vel	Uvu	Pha	Glo
Stop	p b		t̪ d̪	t d	ʈ ɖ			c ɟ	k ɡ	q ɢ		ʔ
Nas	m		n̪	n	ɳ			ɲ	ŋ	ɴ		
Trill				r						ʀ		
Flap		v	ɾ̪	ɾ	ɽ							
Fric	ɸ β	f v	θ ð	s z	ʂ ʐ	ʃ ʒ	ɕ ʑ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
App			r̪	ɹ	ɻ			j	ɰ			
L-App			l̪	l	ɭ			ʎ				

Table 5.11 Analysis of obstruent consonants with a "voiceless" diacritic

Sounds	Language	Examination	Analysis
[b̥ d̥ ɟ̊ g̊]	Ndebele	No [p t c k]	[p t c k]
[ɣ̥]	Breton	Lenis [f]	[f] (vs. [ff])
[ʒ̥]	Georgian	Error	[dz]

Table 5.12 Consonants in the category "others" in Table 5.8

Group	Type	Members
Tense	17	[b̏ c̀ç d̏ ɣ̀ k̀ ɬ̀ m̏ n̏ p̀ s̀ ʃ̀ t̀ t̪̀ ts̀ t̀ʃ x̀ χ̀]
Lax	12	[b̂ d̂ ĝ ĵ k̂ m̂ n̂ t̂ b* p* ʋ* w*]
Glottal	10	[ʔb ʔd ʔj ʔk ʔl ʔm ʔn ʔp ʔt ʔw]
Lateral	9	[bɮ tɬ dl tl l̪θ ɬ ɮ l̟ ɹ]
Breathy	6	[kʷɦ pɦ tɦ tʃɦ t̪s̪ɦ wɦ]
Diacritic	11	[n̞ s̞ d̻ n̻ s̻ z̻ ř š v̌ n̟ r̟]
Remaining	7	[fʃ ʃtʃ βm vɾ s̪v z̪v ɾð]

Voiceless obstruents are normally transcribed with regular IPA symbols, such as [p t k], without a voiceless diacritic. Therefore, we ask whether those with the diacritic can contrast with one without, such as [b̥] vs. [p]. The result is shown in Table 5.11.

In Ndebele, [b̥ d̥ ɟ̊ g̊] do not contrast with unaspirated [p t c k]. Therefore, we can use the latter set of symbols, as other authors of the language do (Pelling 1975; Mawadza 2009). In Breton, [ɣ̥] is the lenis counterpart of fortis [f] (Press 1987: 258), where lenis sounds are short and fortis ones long (Press 1987: 14). We can therefore represent the pair as [f] (lenis) and [ff] (fortis) instead. Finally, [ʒ̥] is is a clerical error for [dz]; the latter is used in the source (Cherchi 1999: 2) and other studies of Georgian (e.g. Aronson 1990; Hewitt 1995). Thus, there is no evidence for a contrast between a devoiced obstruent and a voiceless one.

5.2.3 Other consonants

The category "others" contains seventy-two consonants. They can be divided into several groups, shown in Table 5.12.

Tense consonants are found in four languages, all of which can be analyzed in terms of more common features. This is shown in Table 5.13.

Korean is often reported to have three series of voiceless consonants: aspirated, unaspirated tense, and unaspirated lax. However, it has been shown that, in word-initial positions, the distinction between the latter two does not lie entirely in the consonants themselves but mainly in the following vowel, where a tense onset leads to a high tone and a lax one leads to a low tone (Kim 2000; Kim et al. 2002; Kim and Duanmu 2004). In addition, in medial positions, tense consonants remain voiceless

Table 5.13 Analysis of "tense" consonants in P-base. [k̀] is found in two languages, while the others are found in one language each.

Language	Consonants	Analysis
Korean	[c̀ç k̀ p̀ s̀ t̀]	Voice/aspiration
Godoberi	[ɬ̀ ʃ̀ t̪̀ tś̀ t̀ʃ x̀ χ̀ k̀]	[−aspirated]
Warembori	[b̏ d̏ m̏ n̏]	Long
Mbili	[ɣ̀]	Long

but lax consonants are voiced. Therefore, an alternative analysis is available with regular features: "tense" consonants are voiceless unaspirated, and "lax" consonants are underlyingly voiced. In addition, the tonal difference in word-initial position can be derived from the voicing difference in the onset, similar to the case of standard tonogenesis (Kim and Duanmu 2004).

In Godoberi, consonants can be voiced, voiceless tense, and voiceless lax. According to Kodzasov (1996: 1–6), lax consonants are aspirated and tense ones are not. On the other hand, according to Hewitt (2006: 261), tense (or fortis) consonants are "voiceless unaspirated or geminate." Either way, there is no need to assume the feature "tense," since a regular property, aspiration or length (gemination), is sufficient.

In Warembori (Donohue 1996), "heavy consonants" attract stress, are long, and resist lenition. Therefore, they can again be represented by length (gemination). The same seems true in Mbili, where most consonants have a short–long pair, such as [t t:], [m m:], and [l l:], whereas the two voiced velar fricatives are given as [ɣ ɣ̀], which could be represented as [ɣ ɣ:] instead.

Next we consider lax (or lenis) consonants, which are found in seven languages. Six of them (Auchi, Emhalhe, Ghotuo, Ibilo, Oloma, and Uneme) belong to the Edoid group (of the Niger-Congo family) and are described by the same author (Elugbe 1989). The other language is Maasai (Hollis 1905).

Elugbe (1989: 37–8) states that lenis consonants in the Edoid group are shorter and can alternate with fricatives. In addition, in these sounds "breathy voicing . . . is very much in evidence," and Elugbe indicates it with [h], e.g. writing [b̂ d̂ ĝ ĵ k̂ m̂ n̂ t̂] as [bh dh gh jh kh mh nh th] respectively. Therefore, we can represent the lenis/non-lenis contrast either by length or by breathiness (and in some cases by a fricative/stop contrast).

In the P-base source for Maasai, Hollis (1905: 1–9) (republished in 1971) describes the lax consonants [b* p* v* w*] as being "pronounced in a lazy way by just opening the lips." He lists twenty-six consonants but notes that some "are interchangeable." For example, [gh k] are interchangeable, so are [b* w* ʋ*]. In a later study by Tucker and Mpaayei (1955), for which Hollis wrote the preface, [b* p* v*] are no longer

Table 5.14 Analysis of consonants involving laterals. [dl tɬ] will be discussed in Chapter 7

Sound	Languages	Check	Analysis
[l̟]	17	No [l̪]	[l̪]
[ɬ]	58	No [l̥] or [l̥ʲ]	[l̥] or [l̥ʲ]
[ɮ]	17	No [l̤] or [l̤ʲ]	[l̤] or [l̤ʲ]
[tl]	6	No [tɬ]	[tɬ]
[bɮ]	1	No [bl]	[bl]
[l̪θ]	1	No [θ] or [ð]	[θ] or [ð]
[ɺ]	11	No [l] or [ɾ]	[l] or [ɾ]

included (nor is [gh]). In addition, the contrast between [w w*] is represented as [w wu] instead. Therefore, there seems to be no need for a feature "lax."

Next we consider the glottal consonants [ʔb ʔd ʔj ʔk ʔl ʔm ʔn ʔp ʔt ʔw]. "Glottalization" refers to stiff or partially constricted vocal folds, which is not in itself controversial. What we need to find out is whether there is a contrast between pre-glottalization and post-glottalization, such as [ʔb] vs. [bʔ], or between glottalization and creakiness, such as [ʔb] vs. [b̰]. A search through P-base yields no such contrast.

Next we consider the set [bɮ tɬ dl tl l̪θ ɬ ɮ l̟ ɺ], which involve laterals. I shall argue in Chapter 7 that [dl tɬ] are possible complex sounds (so is [bl]). The analysis of the others in this group is shown in Table 5.14.

[l̟ ɬ ɮ] can be represented with the more common symbols [l̪ l̥ l̤] (or [l̪ l̥ʲ l̤ʲ]) respectively, similar to the case in UPSID. [tl bɮ] can be represented as [tɬ bl], to be discussed in Chapter 7 (along with [dl]). [l̪θ] is reported in Ingessana (Crewe 1975), which does not contrast with [ð] (or [θ]); it can therefore be represented by the latter—a solution adopted by Stirtz (2012: 21) in a more recent and more extensive study. Finally, let us consider [ɺ], which is a lateral flap. We have seen in UPSID that there is no contrast between a lateral flap and a regular lateral, or between a lateral flap and a regular flap. In P-base, [ɺ] is found in eleven languages, six of which do not have a regular lateral [l]; in the other five, there is no [ɾ]. Thus, none of the laterals requires a new feature.

Next we consider breathy consonants [kʷʱ pɦ tɦ tʃɦ t̪s̪ɦ wɦ]. The term "breathy" refers to voiced aspiration, which is not controversial itself. However, we should consider voicing agreement in breathy consonants. For example, in [pɦ], [p] is voiceless while [ɦ] is voiced, which are incompatible features within a sound (Chapter 1). Therefore, we must examine whether [pɦ] contrasts with a more common sound, such as [bɦ] (or [bʱ]). The same applies to other breathy consonants. The result is shown in Table 5.15.

Table 5.15 Analysis of breathy consonants, all of which are found in Ikalanga only

Sound	Check	Analysis
[kʷɦ]	No [gʷɦ]	[gʷɦ]
[pɦ]	No [bɦ]	[bɦ]
[tɦ]	No [dɦ]	[dɦ]
[tʃɦ]	No [dʒɦ]	[dʒɦ]
[t̺s̺ɦ]	No [d̺z̺ɦ]	[d̺z̺ɦ]
[wɦ]	Same voicing	[wɦ]

Of the breathy consonants found in Ikalanga, [wɦ] is not a problem, since both parts are voiced. The other five have a voiceless first part, none of which contrasts with a voiced one. Therefore, all the breathy consonants can be analyzed as underlyingly voiced.

Next we consider the set [n̞ s̞ ř š v̌ d̻ n̻ s̻ z̻ n̟ r̟], each of which involves a special diacritic. We again examine whether each can contrast with a more common sound. The result is shown in Table 5.16.

[n̞ s̞] are found in Ndebele. In the source (Ziervogel 1959: 21), these sounds are described as "post-alveolar," a place that lies between alveolar and palatal. A more common feature for this place is retroflex, or [ɳ ʂ], which Ndebele does not otherwise have. The diacritic [ˇ] has two meanings; in the IPA system it means a short sound, whereas in popular usage it means a palatal sound. [ř] is found in Czech, which does not contrast with a palatal [r], or [rʲ]. [š] is reported in Nuer, which does not contrast

Table 5.16 Analysis of consonants involving a diacritic. A question mark in the diacritic column indicates that the IPA meaning of the diacritic does not match the description of the sound

Sound	Count	Diacritic	Source	Analysis	Language
[n̞]	1	?	Post-alveolar	[ɳ]	Ndebele
[s̞]	1	?	Post-alveolar	[ʂ]	Ndebele
[ř]	1	Palatal	'soft"	[rʲ]	Czech
[š]	1	Palatal	No [ʃ] or [s]	[ʃ] or [s]	Nuer
[v̌]	1	Short	Flap	[ʋ]	Higi
[d̻]	2	Laminal	No [dʲ]	[dʲ]	Hixkeryan, Dyerbal
[n̻]	2	Laminal	No [nʲ]	[nʲ]	Dyerbal, kamba
[s̻]	3	Laminal	No [sʲ]	[sʲ]	Kannada, Arabic, Basque
[z̻]	1	Laminal	No [zʲ]	[zʲ]	Arabic
[n̟]	17	Advanced	No [n̺]	[n̺]	Irish dialects
[r̟]	17	Advanced	No [r̺]	[r̺]	Irish dialects

Table 5.17 Remaining consonants in the "others" category

Sound	Count	Language	Source	Analysis
[fʃ]	1	Sotho	Cluster	[f]+[ʃ]
[ʃtʃ]	1	Russian	Cluster	[ʃ] + [tʃ]
[βm]	1	Waffa	[β̃]	[β̃] or [w̃]
[vɾ]	1	Margi	Flap, marginal	marginal
[s̪v]	2	Karanga (2 dialects)	[sʷ]	[sʷ]
[z̪v]	2	Karanga (2 dialects)	[zʷ]	[zʷ]
[ɾð]	1	Tacana	Laminal flap	[ɾʲ]

with [ʃ] (a palatal [s]) or a regular [s]. [v̌] is reported in Higi. The source describes it as a flapped fricative (Mohrlang 1972: 23), for which there is a more common symbol, [ⱱ], which the language does not use. The term "laminal" is often used to refer to palatal (or palatalized) consonants. Therefore, we check whether [d̻ n̻ s̻ z̻] contrast with [dʲ nʲ sʲ zʲ] and the answer is no. Therefore, we can represent [d̻ n̻ s̻ z̻] as [dʲ nʲ sʲ zʲ]. Finally, [n̟] (a nasal) and [ɾ̟] (a tap) are reported in Irish (17 dialects). The diacritic "advanced" refers to place of closure more forward than what is indicated by the IPA symbol. In the case of [n̟ ɾ̟], the closure is more forward than alveolar, which is dental. Therefore, we ask whether [n̟ ɾ̟] contrast with [n̪ ɾ̪] and the answer is no. Therefore, we can represent [n̟ ɾ̟] as [n̪ ɾ̪].

The last set of sounds to consider consists of [fʃ ʃtʃ βm vɾ s̪v z̪v ɾð]. They are analyzed in Table 5.17.

[fʃ] is found in Sotho. According to the source (Doke and Mofokeng 1957: 23), [fʃ] only occurs in a specific context, where a verb ends in [f] and the passive suffix starts with [ʃ]. In addition, [fʃ] "is merely alternative to" [ʃ]. Therefore, there is no reason to consider [fʃ] to be a phoneme but a cluster of two sounds.

[ʃtʃ] is a clerical error. It is the pronunciation of the Russian orthographic letter "щ," not of a phoneme (Unbegaun 1957). In phonemic studies of Russian (such as Hamilton 1980), [ʃtʃ] is not included in the phoneme inventory. It is of interest, too, that [ʃtʃ] alternates with [ʃʃ] and can be represented by two Russian letters, "сч" or "зч" (Unbegaun 1957: 7).

[βm] is reported in Waffa. In the source (Stringer and Hotz 1973: 526), it is described as "a nasalized voiced bilabial fricative." If so, the representation should be [β̃]. Such a sound, while rare, has compatible features, because [β] and nasalization involve independent articulatory gestures. Alternatively, we can represent the sound as [w̃] (which Waffa does not otherwise have), which is found in nineteen other inventories in P-base.

[vɾ] is reported in Margi. In the source (Hoffman 1963: 25), it is a "voiced labiodental flap," whose IPA symbol is [ⱱ]. It is worth noting, too, that [vɾ] is marginal in Margi, since it is only found "in ideophones only" (Hoffman 1963: 25).

TABLE 5.18 Basic consonants in UPSID and P-base. Two sounds, [ɕ ʑ], are found in P-base only. Twenty-nine sounds, [b̪ t^{PA} d^{PA} m̥ m̪ n̪̊ n̥ ɳ̊ ɲ̊ ɲ ȵ̊ ŋ̊ r̪ r̥ r^{Ret} r^{Pal} ɾ^{PA} ɸ^{App} β^{App} ʋ ɹ̥ j^{PA} j̊ ʁ^{App} l̪̊ l̥ ɭ̊ l^{PA} ʟ̥ L], are found in UPSID only, most of which are voiceless sonorants. The remaining fifty-nine sounds are found in both databases

	BL	LD	ID	D/A	Ret	PA	PPal	Pal	Vel	Uv	Pha	Gl
Plo	p b	b̪	t̪ d̪	t d	ʈ ɖ	t^{PA} d^{PA}		c ɟ	k g	q ɢ		ʔ
Nas	m̥ m	ɱ	n̪̊ n̪	n̥ n	ɳ̊ ɳ	ɲ̊ ɲ		ȵ̊ ȵ	ŋ̊ ŋ	ɴ		
Trill			r̪	r̥ r	r^{Ret}			r^{Pal}		ʀ		
Fl/T		ⱱ	ɾ̪	ɾ	ɽ	ɾ^{PA}						
Fric	ɸ β	f̪ v̪	θ ð	s z	ʂ ʐ	ʃ ʒ	ɕ ʑ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Ap	ɸ^{Ap} β^{Ap}	ʋ	ɾ̪	ɹ̥ ɹ	ɻ	j^{PA}		j̊ j	ɰ	ʁ^{App}		
L-Ap			l̪̊ l̪	l̥ l	ɭ̊ ɭ	l^{PA}		ʎ	ʟ̥ ʟ			

[s̪v] and [z̪v] are reported in two varieties of Karanga. According to the source (Marconnès 1931: 7), these sounds are phonetically [s̪ʷ] and [z̪ʷ] (with bilabial rounding) and should be so transcribed. They are represented orthographically as "sv" and "zv" in order to distinguish them from [sw] and [zw], which are clusters of two sounds each.

Finally, [ɾð] is reported in Tacana as a flap. According to the source (Key 1968: 28), Tacana has two flaps, one involving the "tongue tip" and one the "tongue blade"—a contrast also used to distinguish [s] and [ʃ]. The contrast could be interpreted as either one of palatalization [ɾ]-[ɾʲ], or one of retroflection [ɾ]-[ɽ].

5.3 Summary

Our analysis of UPSID and P-base shows that there are fewer basic consonants than previously assumed. In addition, the two databases yield a similar inventory of basic consonants, seen in Table 5.7 for UPSID and Table 5.10 for P-base, which we merge to yield Table 5.18. The places of articulation are bilabial (BL), labio-dental (LD), interdental (ID), dental or alveolar (D/A), retroflex (Ret), palatal-alveolar (PA), pre-palatal (PPal), palatal (Pal), velar (Vel), uvular (Uv), pharyngeal (Pha), and glottal (Gl). The manners of articulation are plosives (Pl), nasal (Nas), trill, flap or tap (Fl/T), fricative (Fric), approximant (Ap), and lateral approximant (LAp).

The table still has many empty cells, and several issues could be further explored. For example, do flaps/taps contrast with stops? Do trills contrast with fricatives? Do trills contrast with flaps/taps? Preliminary examinations suggest that some further reduction is possible. For example, no language in P-base has a contrast between the trill [ʀ] and the fricative [ʁ]. However, we leave these questions open for now.

6

A feature system

Many feature systems have been proposed before and there is a large set of feature terms with overlapping meanings in the literature. For example, "voiceless" is similar to the binary feature value "[–voice]" or "[+stiff] vocal folds," and "oral" is similar to "[–nasal]," or "[+raised] velum." In the preceding chapters I have used fairly common feature terms to refer to contrasts. Since we have found fewer contrasts than previously assumed, there is the question of which subset of traditional features to keep. In particular, some features focus on acoustic effects, such as "voiceless," and some focus on articulatory gestures, such as "[+stiff] vocal folds" (Halle and Stevens 1971). Should we focus on articulatory gestures, acoustic properties, or auditory impressions? In addition, can features refer to abstract concepts, devoid of phonetic correlates? Moreover, should features be binary, multi-valued, or mono-valued?

In this chapter we consider such questions. I start by a review of several major proposals. Then, I outline a system in which features are gestures of active articulators. Next I discuss the representation of vowels, consonants, and tones, followed by a discussion of feature specification. Finally, I discuss variations in the phonetic realization of features.

6.1 Previous proposals

In this section, I review several feature systems that have been proposed before: Jakobson et al. (1952), Chomsky and Halle (1968), Clements (1985), Clements and Hume (1995), Browman and Goldstein (1989), Halle (1995; 2003), and Ladefoged (2007).

6.1.1 Jakobson et al. (1952)

Jakobson et al. (1952) propose a system of fifteen binary features, shown in Table 6.1.

Some feature terms are not commonly used in other systems, and for them a more common term (or terms) is given in the last column. For example, "interrupted" is similar to a "stop" consonant and "checked" is similar to a glottalized consonant or vowel.

A Theory of Phonological Features. First edition. San Duanmu.

The feature types "source," "envelope," and "resonance" imply a source–filter model of speech production. Specifically, Jakobson et al. (1952) assume the function W = T • S, where W is the speech signal (waveform), T the transfer function (the vocal tract), and S the sound source. Naturally, there is a preference to define features in acoustic terms. For example, "resonance" features are defined by formant frequencies, formant positions, formant movement, or the energy level, as indicated in the last column in Table 6.1, regardless of whether the sounds under each feature share any articulatory similarities. For example, "grave" covers dorsal (tongue body) consonants [k g], labial consonants [p b], and back vowels [ɑ u ɯ], but not coronal (tongue tip) consonants [t d] or front vowels [æ i y]. This means

Table 6.1 A system of fifteen binary features proposed by Jakobson et al. (1952). The "source," "envelope," and "resonance" features are also called "inherent" features

Feature	Type	Example	Notes
Vocalic/non-vocalic	Source	[a r]/[p j]	
Consonantal/non-consonantal	Source	[r p]/[a j]	
Voiced/unvoiced	Source		
Interrupted/continuant	Envelope		Stop
Checked/unchecked	Envelope		Glottalized
Strident/mellow	Envelope	[s]/[θ]	
Compact/diffuse	Resonance	[k c]/[p t] [ɑ æ]/[u i]	F1 higher in compact; dorsal/non-dorsal low/high
Grave/acute	Resonance	[k p]/[c t] [ɑ u]/[æ i]	F2 is closer to F1 in grave; non-coronal/coronal back/front
Flat/plain	Resonance	[y]/[i]	Downward shift of formants in flat; round/non-round
Sharp/plain	Resonance	[mʲ]/[m]	Rise in F2 in sharp; palatalization
Tense/lax	Resonance	[s p]/[z b] [i]/[ɪ]	Longer duration and more energy in tense
Nasal/oral	Resonance		
Syllabic/non-syllabic	Prosodic	[i]/[j]	
Stressed/unstressed	Prosodic		
Long/short	Prosodic		

that consonants that use the lips are always grave, but vowels that use the lips (for rounding) may or may not be grave. It seems that the feature system is guided by a goal of minimizing the total number of features, with little regard to articulatory similarities or differences.

Nevertheless, the feature system is not minimal, in that the "prosodic" features are redundant: the distinctions they make can be represented by syllable structure, which is independently needed (Chapter 1). On the other hand, the feature system is insufficient, since it only covers pulmonic sounds (those using lung air). Moreover, the feature system raises a theoretical question: Are acoustically defined features convenient technical tools for processing speech by machine, or do they represent real mental elements that humans use to produce and understand speech? Finally, the feature system is not verified against a database of phonemic contrasts or one of sound classes.

6.1.2 Chomsky and Halle (1968)

A significantly revised and expanded feature system is offered in Chomsky and Halle (1968: 299–300), which includes thirty-three binary features, shown in Table 6.2. Following a common practice, binary features are shown in square brackets.

A main claim of Chomsky and Halle (1968), not made in Jakobson et al. (1952), is that features are part of universal grammar, the innate knowledge for language. Specifically, features belong to "substantive universals," which are innate categories common to all languages (Chomsky and Halle 1968: 4). In addition, while Jakobson et al. (1952) favor acoustic definitions of features, Chomsky and Halle (1968) favor articulatory definitions—perhaps because it is more natural to assume that humans know how to make speech sounds than to assume that they know the acoustic properties of speech sounds, such as formants and amplitude.

Since they share a focus on articulation, the feature system of Chomsky and Halle (1968) shows many similarities to the IPA system (International Phonetic Association 1999). In particular, the feature types "cavity," "manner," and "source" of Chomsky and Halle (1968) correspond closely to "place," "manner," and "voicing" in the IPA system. In addition, most "prosodic" features (including "pitch features") and the "major class" features of Chomsky and Halle (1968) are found in the IPA system, too. On the other hand, the feature system of Chomsky and Halle (1968) does not always explicitly refer to articulation. For example, there is no reference to the "labial" articulator. Instead, a labial consonant is referred to by [–coronal] (closure is not made by the tongue tip) and [+anterior] (closure is in front of the alveolar ridge). One has to deduce indirectly that a closure made in front of the alveolar ridge and not by the tongue tip has to be "labial." Similarly, it is not explicitly stated what part of the tongue the "cavity" features [high],

Table 6.2 A system of thirty-three binary features proposed by Chomsky and Halle (1968: 299–300). The last eight features (26–33) are, however, little discussed

	Feature	Type
1	[sonorant]	Major class
2	[vocalic]	Major class
3	[consonantal]	Major class
4	[coronal]	Cavity
5	[anterior]	Cavity
6	[high]	Cavity
7	[low]	Cavity
8	[back]	Cavity
9	[round]	Cavity
10	[distributed]	Cavity
11	[covered]	Cavity
12	[glottal constrictions]	Cavity
13	[nasal]	Cavity
14	[lateral]	Cavity
15	[continuant]	Manner
16	[primary release]	Manner
17	[secondary release]	Manner
18	[velaric suction (clicks)]	Manner
19	[implosion]	Manner
20	[velaric pressure]	Manner
21	[ejectives]	Manner
22	[tense]	Manner
23	[heightened subglottal pressure]	Source
24	[voice]	Source
25	[strident]	Source
26	[stress]	Prosodic
27	[high]	Pitch
28	[low]	Pitch
29	[elevated]	Pitch
30	[rising]	Pitch
31	[falling]	Pitch
32	[concave]	Pitch
33	[length]	Prosodic

[low], and [back] refer to, and this leaves room for questions. For example, Walker (1993), Rice (1995), and Clements and Hume (1995) propose that front vowels involved the tongue tip (coronal), and Halle and Stevens (1969) and Pulleyblank (1997) propose that low vowels involved the tongue root (radical).

Another point of interest is that some feature terms have similar meanings. For example, [+anterior] (in front of the alveolar ridge), [–back] (tongue body is fronted), and [–covered] (tongue root is fronted) all share the meaning "front." Evidently, this is because they refer to different articulators. On the other hand, [high] and [low] (of the tongue) and [high] and [low] (of pitch) are not distinguished, presumably because they are not often used together and are less likely to cause confusion.

Finally, it is worth noting that, like the feature system of Jakobson et al. (1952), the feature system of Chomsky and Halle (1968) is not verified against a database of phonemic contrasts or one of sound classes. What is clear, though, is that some features are excessive. For example, [vocalic] and [length] can be represented by syllable structure (Chapter 1). Similarly, it is unlikely that we need six binary features to represent pitch (Duanmu 1996; Yip 2002).

6.1.3 Clements (1985)

Clements (1985) proposes that features should be organized into a tree structure, which he calls "feature geometry." The proposal is shown in Table 6.3, where a higher node dominates two or more nodes below (to its right here). For example, the "manner" node dominates five features, the "supra-laryngeal" node dominates "manner" and "place," and the "root" node dominates "laryngeal" and "supra-laryngeal."

Justification of higher nodes in the tree structure is based on the degree of assimilation in sound change. Specifically, each node represents a possible point of sound change. When a change occurs at a "terminal node," it affects one feature, such as [voiced]. When a change occurs at a higher node, all features dominated by it will change. For example, a change at

Table 6.3 A feature system proposed by Clements (1985)

Higher nodes			Features
Root	Laryngeal		[spread]
			[constricted]
			[voiced]
			(tone)
	Supra-laryngeal	Manner	[nasal]
			[continuant]
			[strident]
			[sonorant]
			[lateral]
			[consonantal]
		Place	[coronal]
			[anterior]
			[distributed]
			[round]
			[high]
			[back]
			…

the "laryngeal" node will change all voicing and aspiration features, but not other features. Finally, if the change occurs at the "root" node, it is a complete change of one sound to another, as in AB → BB, where A and B are two sounds and A has changed to B.

It can be argued that the organization of features is independent of what features we need. The former is based on patterns of sound change and sound classes, which can sometimes be problematic (Mielke 2008). The latter is based on contrasts, which must be recognized regardless of how sounds change or how they form classes. The features of Clements (1985) are mostly those of Chomsky and Halle (1968). In addition, the higher organization nodes "place," "manner," and "laryngeal" are similar to those in the IPA system. The proposal is preliminary, though, and its feature list is rather short, compared with that of Chomsky and Halle (1968). For example, tone features (shown in parentheses) are not discussed. Similarly, not all place features are discussed. And there is no discussion of non-pulmonic sounds, such as clicks and implosives. Therefore, as far as representing contrasts is concerned, Clements (1985) offers little new.

6.1.4 Clements and Hume (1995)

A more detailed feature tree is offered by Clements and Hume (1995) and shown in Table 6.4. As in Clements (1985), arguments for the higher nodes are based on assimilation processes.

TABLE 6.4 A feature system proposed by Clements and Hume (1995)

Higher nodes					Features	
Root					[sonorant]	
					[approximant]	
					[vocoid]	
					[nasal]	
				Laryngeal	[spread]	
					[constricted]	
					[voice]	
	Oral				[continuant]	
		C-place			[labial]	
					[dorsal]	
					[coronal]	[anterior]
						[distributed]
			Vocalic	Aperture	[open]	
				V-place	[labial]	
					[dorsal]	
					[coronal]	[anterior]
						[distributed]

The main point of interest is a split between consonant representation and vowel representation. Consonants lack the “vocalic” node. Vowels have the “vocalic” node, and the features below it, but no other features under the “C-place” node. Thus, the features [labial], [dorsal], [coronal], [anterior], and [distributed] show up twice in the tree, once directly under the C-place node (for consonants) and once under the V-place node for vowels. In addition, vowels have the extra “aperture” feature [open]. Consonants with a secondary articulation will have features both directly under the C-place node and under the V-place node. For example, [k^w] has [dorsal] directly under C-place and [labial] under V-place, and [b^w] has two [labial] features, one directly under C-place and one under V-place.

There are some other points of interest. First, features have different numbers of values. Some are one-valued (or “privative”), such as [labial], [coronal], and [dorsal], which are either present or absent. Some features have binary values, such as [sonorant]. Some features have multiple values, such as [open] (for vowel height). Second, the traditional vowel features [back] and [round] are replaced by [labial], [coronal], and [dorsal], so that [+round] is the presence of [labial], [–round] is the absence of [labial], [–back] is the presence of [coronal] (and the absence of [dorsal]), and [+back] is the absence of [coronal] and the presence of [dorsal]. Third, the notion of active articulators is given little consideration. For example, it is not stated what articulator performs the feature [nasal]. Similarly, [labial] is used under C-place as an articulator, which can perform the feature [continuant], and as a feature under V-place for what is traditionally called “round.” Finally, as far as representing contrasts is concerned, the system has little to say, since it is not checked against a phoneme inventory database.

6.1.5 Browman and Goldstein (1989)

Browman and Goldstein (1989) propose that features are gestures of articulators—an idea that is absent in Clements (1985) and Clements and Hume (1995). The proposal follows from the motor theory of speech perception (Liberman et al. 1967; Liberman and Mattingly 1985), according to which speech sounds are perceived in terms of articulatory gestures that produce them. The proposal is shown in Table 6.5.

In many studies (e.g. Ladefoged and Johnson 2011), an “articulator” can be a static location in the vocal tract, such as the hard palate, or a moveable part, such as the tongue body. In Browman and Goldstein (1989), the term refers to the latter only, and there are six articulators: the lips, the tongue tip, the tongue body, the tongue root, the velum, and the glottis. All articulators except the tongue root have the feature CD (constriction degree), which normally has five values: “closed” (for stops), “critical” (for fricatives), “narrow” (for approximants and high vowels), “mid” (for mid vowels), and “wide” (for low vowels), although CD may have just two values (open or closed) for the velum.

The feature CL (constriction location) is similar to the IPA notion of “the place of articulation” and has eleven values, which are protruded (lips), labial, dental, alveolar, post-alveolar, palatal, velar, uvular, pharyngeal, and glottal (Browman and Goldstein 1989: 227). Each articulator only has a subset of the CL values. For example, the lips has

Table 6.5 A feature system proposed by Browman and Goldstein (1989: 223) (CD = constriction degree; CL = constriction location; CS = constriction shape)

Higher nodes			Articulators	Gestures
Vocal tract	Oral		Lips	CD
				CL
				stiffness
		Tongue	Tip	CD
				CL
				CS
				stiffness
			Body	CD
				CL
				CS
				stiffness
			Root	CL
				stiffness
			Velum	CD
				stiffness
			Glottis	CD
				CL
				stiffness

the CL values protruded, labial, and dental only, and the tongue tip has the CL values of labial, dental, alveolar, post-alveolar, and palatal only. Some CL values are shared by two adjacent articulators. For example, both the lips and the tongue tip can have the CL values "labial" and "dental," and both the tongue tip and the tongue body can have the CL value "palatal." There is a difference between the two cases, though. In the case of "dental," either the lower lip can move to the upper teeth (to make a "labio-dental") or the tongue tip can move between the teeth (to make an "interdental"), but not both at the same time. Similarly, in the case of "labial," either the lower lip can move to the upper lip (to make a "bilabial") or the tongue tip can move to the upper lip (to make a "linguo-labial"), but not both at the same time. However, in the case of "palatal," both the tongue tip and the tongue body are involved at the same time (Browman and Goldstein 1989: 225).

The feature CS (constriction shape) mainly concerns the shape of the tongue—in particular whether the constriction involves the tongue tip (traditionally called "apical") or the tongue blade (traditionally called "lamina" or "distributed"), and whether the tongue is "narrow" (for laterals, such as [l]) or not (for non-laterals). Finally, the feature "stiffness" is the most tentative; it is intended to cover a variety of things, such as the difference between glides and high vowels (e.g. [j] is more stiff than [i]), the effect of "stress and speech rate," and possibly "pitch control."

Several higher nodes above the articulators have been added in order to create a tree structure, which is meant to resemble the feature geometry of Clements (1985). However, not much discussion is offered to justify the higher nodes.

The feature system is not checked against a comprehensive database, and there are parts that seem to be missing and parts that seem to be excessive. For example, the articulator "larynx" should be added, in order to represent implosives and ejectives. Similarly, tone features need to be specified. On the other hand, it is questionable whether we need five values for CD (constriction degree), given the fact that there are only two degrees of vowel height (Chapter 4), rather than three. Similarly, it is unclear whether every articulator has a contrastive gesture in "stiffness."

6.1.6 Halle (1995; 2003)

Halle (1995; 2003) proposes a system of six articulators and nineteen binary features. They are shown in Table 6.6. Similar proposals have been offered by Sagey (1986), Ladefoged and Halle (1988), Halle et al. (2000), and Halle (2005).

Like Browman and Goldstein (1989), Halle considers features to be gestures of active articulators, i.e. those that are moveable in the vocal tract (as opposed to static locations, such as the teeth, the alveolar ridge, or the hard palate, which have been

TABLE 6.6 A system of six articulators and nineteen binary features, proposed by Halle (1995; 2003)

Features	Articulators	Other terms
[consonantal]	(various)	
[sonorant]	(various)	
[suction]	(various)	
[continuant]	(various)	
[strident]	(various)	
[lateral]	(various)	
[nasal]	Velum	Soft palate
[stiff vocal folds]	Glottis	Larynx; Laryngeal; Vocal folds
[slack vocal folds]	Glottis	Larynx; Laryngeal; Vocal folds
[spread glottis]	Glottis	Larynx; Laryngeal; Vocal folds
[constricted glottis]	Glottis	Larynx; Laryngeal; Vocal folds
[retracted tongue root]	Tongue root	Radical
[advanced tongue root]	Tongue root	Radical
[round]	Lips	Labial
[anterior]	Tongue tip	Coronal; Tongue blade
[distributed]	Tongue tip	Coronal; Tongue blade
[high]	Tongue body	Dorsal
[low]	Tongue body	Dorsal
[back]	Tongue body	Dorsal

called "passive articulators"). Indeed, Ladefoged and Halle (1988: 582) consider the recognition of active articulators to be an issue of "paramount importance."

I have used the same articulator names as those in Browman and Goldstein (1989). Common alternative articulator names are shown in the right-hand column. Specifically, "velum" is called "soft palate" in Halle (1995; 2003); "glottis" is called "larynx" in Halle (1995), "laryngeal" in Ladefoged and Halle (1988), and "vocal folds" in Halle (2005); "tongue root" is called "radical" in Ladefoged and Halle (1988); "lips" is called "labial" in Halle (1995); "tongue tip" is called "coronal" in Halle (1995) and "tongue blade" in Halle (2003); and "tongue body" is called "dorsal" in Halle (1995).

Thirteen of the features are performed by a specific articulator; for example, [nasal] is performed by the velum only. The other six features can be performed by two or more articulators; for example, [–continuant] can be performed by the lips in [p], the tongue tip in [t], and the tongue body in [k].

The proposal of Halle (1995; 2003) is not checked against a phoneme inventory database and some parts seem to be missing. For example, the articulator "larynx" should be added, in order to represent implosives (which require the lowering of the larynx) and ejectives (which require the raising of the larynx). Similarly, we may need another feature for the lips in order to distinguish bilabial [ɸ] and labiodental [f]. On the other hand, some parts of the feature system seem to be excessive. For example, we may not need both [high] and [low] for the tongue body, given the fact that there are only two degrees of vowel height (Chapter 4), rather than three. Similarly, the features [sonorant] and [consonantal] may be redundant, and we may not need both [advanced] and [retracted] for the tongue root.

A couple of other remarks can be made. First, the feature [suction] is used to describe clicks, but its articulator is not very clear. Indeed, it is left open whether [suction] is a gesture of any of the articulators. Second, as in Chomsky and Halle (1968), some feature terms have similar meanings. For example, [+anterior] (of the tongue tip), [–back] (of the tongue body), and [+advanced] (of the tongue root) share the meaning "front." If these features are dominated by different articulators, there might not be a need to use different terms.

6.1.7 Ladefoged (2007)

Ladefoged (2007: 162) proposes a system of twenty-three features, shown in Table 6.7. A similar system proposed in Ladefoged (1971: 92–4) contains twenty-seven features, some no longer used, such as [accent] and [gravity].

Each feature may have one or more values, which Ladefoged places in brackets. For example, "labial" has the values [bilabial] and [labiodental], "laminal" has the values [linguo-labial], [interdental], [laminal-dental], [laminal-alveolar], and [laminal-postalveolar], and "sub-apical" has the single value of [sub-apical palatal].

Table 6.7 A system of twenty-three features and seven articulators, proposed by Ladefoged (2007: 162)

Class	Articulator	Feature	Values
Place	Labial	[labial]	2
	Coronal	[laminal]	5
		[apical]	3
		[sub-apical]	1
		[tongue shape]	4
	Dorsal	[dorsal]	3
	Radical	[radical]	2
	Glottal	[laryngeal]	1
Manner		[trill]	2
		[stricture]	3
	Dorsal	[height]	5
		[backness]	3
	Labial	[rounding]	3
		[lipprotrusion]	2
	Radical	[tongue root]	2
Lateral	Coronal	[laterality]	2
Nasal	Velum	[nasal]	2
Glottal	Glottal	[glottal stricture]	5
		[glottal timing]	1
	Larynx	[glottal movement]	2
Airstream		[pulmonic]	2
		[velaric]	1
Syllable		[syllabicity]	2

Similarly, "glottal stricture" has the values [voiceless], [breathy voice], [voiced], [creaky voice], and [closed], and "glottal timing" has the single value of [aspirated].

For most features, the articulator name is given by Ladefoged, such as "labial," "coronal," "dorsal," "glottal," and "radical." For features without a specified articulator, I have added one whenever reasonably possible; for example, I added "velum" for the feature [nasal], "coronal" for the feature [lateral], and "larynx" for the feature [glottal movement]. There are altogether seven articulators: the lips (labial), the tongue tip (coronal), the tongue body (dorsal), the tongue root (radical), the velum, the glottis (glottal), and the larynx. These (except the larynx, whose feature [glottal movement] has the values [ejective] and [implosive]) are similar to the articulators proposed in Browman and Goldstein (1989) and Halle (1995; 2003).

The feature system is similar to that of the IPA, where features are defined in articulatory terms and grouped into several classes according to the "place" of articulation, the "manner" of articulation, the glottal properties, and airstream mechanisms. The feature system is not explicitly checked against a phoneme inventory database; but given that the author is a leading scholar in the study of consonants and vowels and has worked on many languages around the world, the feature system is sufficient to represent all known contrasts.

However, the feature system seems excessive. First, some features are redundant. For example, "syllabicity" is meant to distinguish high vowels from glides, such as [u] vs. [w], which can be distinguished by syllable structure. Similarly, "pulmonic" is meant to distinguish [fortis] and [model] consonants, such as [t*] vs. [t] in Korean, but the contrast has been questioned (Kim et al. 2002; Kim and Duanmu 2004). Second, some feature values seem to be excessive. For example, "laminal" has the values [linguo-labial], [interdental], [laminal-dental], [laminal-alveolar], and [laminal-postalveolar], but according to Ladefoged and Maddieson (1996: 40), [linguo-labial] and [interdental] never contrast in any language. Therefore, we can omit [linguo-labial]. There is some overlap between features, too, such as "labial" and "round." Finally, the distinction between articulators and features is not strictly made; for example, "labial" is both an articulator (for "round") and a feature.

6.1.8 Summary

We have seen that, in representing contrasts, current feature systems offer either too few distinctions or too many. In addition, features have been defined in various ways. In particular, features have been defined in acoustic terms, articulatory terms, auditory terms, or a mixture of them. Features have also been defined in abstract terms. For example, Fudge (1967) argues that features should be abstract. Ladefoged and Johnson (2011) consider "tense" be an abstract feature in English, devoid of phonetic content. Jakobson et al. (1952) and Halle (1983) consider features to be abstract mental categories that mediate between articulation and acoustics. Hale and Reiss (2000) consider all features to be completely abstract mental concepts.

With regard to the number of values a feature can have, there are also different views. Some believe that all features have binary values, represented with a plus or minus sign (e.g. Jacobson et al. 1952; Halle 1957; Chomsky and Halle 1968). In a slightly different view, a feature can have three equipollent values, such as [+voice], [–voice], and [ovoice] (Vaux 2009). Some believe that a feature is a phonetic scale, which can take on many values (Ladefoged 1972). In yet another view, all features are "privative" (single-valued), where the presence or absence of a feature offers a two-way contrast (e.g. Rice 1992). For example, the presence of [nasal] corresponds to [+nasal] of the binary representation, and the absence of [nasal] corresponds to [–nasal]. Trubetzkoy (1969) offers yet another view, according to which some features can be binary, some multi-valued, and some mono-valued. Similarly, in Halle (2003), most features are binary, such as [nasal] and [back], but features referring to articulators (such as Labial, Coronal, and Dorsal) are mono-valued. It is worth noting that a multi-valued system can be converted to a binary system by adding more features; for example, a three-height system can be represented with two binary features [high] and [low], where [+high, +low] is excluded. Similarly, a binary system can be converted to a mono-valued system if we interpret the absence of a feature to be its negative value. Moreover, a mono-valued system can achieve multi-valued effects. For example, in Parkinson (1996), the privative feature [closed] can be repeated many times, each repetition indicating a slightly higher vowel.

6.2 A new feature system

In this section I propose a feature system that is minimally sufficient to represent all contrasts. In Chapters 4–6 we have seen that there are fewer contrasts in the world's languages than previously assumed. Therefore, we only need a subset of the features that have been proposed. I choose a set of least controversial features that are defined in articulatory terms. The articulatory approach agrees with the motor theory of speech perception, according to which speech sounds are perceived in terms of the articulatory gestures with which they are pronounced (Liberman et al. 1967; Liberman and Mattingly 1985). The number of values for each feature is determined by contrasts empirically observed in UPSID and P-base, rather than by theoretical assumptions.

6.2.1 Articulators and gestures

I follow a common approach (e.g. Abercrombie 1967; Liberman et al. 1967; Liberman and Mattingly 1985; Sagey 1986; Browman and Goldstein 1989; Ladefoged and Halle 1988; Halle 1995; 2003) and consider features to be articulatory gestures. This is essentially the approach of the IPA system (International Phonetic Association 1999), too, although the IPA does not always explicitly indicate the active articulator of a feature.

The articulator-based definition of features is given in (1), and a set of seven articulators, along with their features, are shown in (2). The Glottis feature [H] is a cover term for a "high" tone, which probably involves stretching and thinning the vocal folds. The articulator Larynx is used for ejectives and implosives, to be discussed in Chapter 7.

(1) Articulator-based definition of features
A feature is a gesture of an articulator, where an articulator is a moveable part in the vocal tract whose gesture(s) can distinguish sounds.

(2) Articulators and their gestures (features)

Articulator	"Manner" features	Other features
Lips (lower lip, labial)	[stop], [fricative]	[location], [round]
Tip (of the tongue, coronal)	[stop], [fricative]	[location], [narrow]
Body (of the tongue, dorsal)	[stop], [fricative]	[high], [location]
Root (of the tongue)		[advanced]
Velum (soft-palate)	[stop]	
Glottis (vocal folds)	[stop]	[stiff], [spread], [H]
Larynx		[raised]

The articulators in (2) are similar to those proposed by Ladefoged and Halle (1988), Browman and Goldstein (1989), Halle (1995; 2003), and Ladefoged (2007). In addition, following Browman and Goldstein (1989), I use a cover feature [location] to represent the forward or backward movement along the vocal tract, which yields traditional "place" features, to be explained below. Moreover, I follow Ladefoged (1980: 492) and Browman and Goldstein (1989) and assume that lateral consonants are made with narrowed tongue tip, hence the feature [narrow].

The manner features [stop] and [fricative] are not available for all articulators, because some articulators lack the full range of gestures. In particular, the velum cannot make a fricative (by forming a narrow opening to the nasal cavity where noise is created). In addition, the tongue root cannot make a full closure (a [stop] gesture). Moreover, the larynx cannot make a stop or fricative. Finally, the glottis can make a stop, but not a fricative or an affricate. Our feature system offers a minimally sufficient set of representations of the sounds created by each articulator. For example, without the feature [fricative], we can still represent what are traditionally called "pharyngeal fricatives" and "glottal fricative." This is shown in (3).

(3) Representing "pharyngeal fricatives" and "glottal fricative" without [fricative]

	Root	Glottis
Pharyngeal fricative	[–advanced]	
Glottal fricative		[+spread]

The feature list in the present proposal is shorter than that of Halle (2003). In particular, on the basis of our analysis of UPSID and P-base, there is no need for [constricted glottis], [slack vocal folds], [constricted pharynx], [distributed], or [low]. In addition, the manner features [consonantal], [sonorant], [continuant], and [strident] are replaced by combinations of [stop] and [fricative]. Nevertheless, we shall see that the present feature system is sufficient for distinguishing all contrasts in UPSID and P-base.

6.2.2 *Interpreting manner features*

In the IPA system, features are divided into two major classes: place and manner. In the present analysis, there are two manner features only, [stop] and [fricative], which can be performed by different articulators. The two features combine to yield four traditional manner features, shown in (4). The analysis is similar to that of Ladefoged (1971), Sagey (1986), Lombardi (1990), and Padgett (1995).

(4) Interpreting "manner" features

Present	Traditional
[+stop, –fricative]	Stop
[–stop, +fricative]	Fricative
[+stop, +fricative]	Affricate
[–stop, –fricative]	Approximant

Among the combinations, [+stop, +fricative] seem problematic. [+stop] requires a complete closure in the vocal tract, whereas [+fricative] requires there to be an opening. If the two gestures are made simultaneously, there seems to be a contradiction, because the vocal tract cannot be both closed and open at the same time. If the two gestures are made in sequence, then we run into another problem: If feature can occur in sequence within a sound, then we predict far more possible sounds than there is evidence for. I shall return to this problem in Chapter 7.

In our analysis, [stop] is applicable to Velum as well, which replaces the traditional feature [nasal]. This is shown in (5).

(5) Interpreting actions of Velum

Present	Traditional	Comment
Velum-[+stop]	Oral	Stop for the nasal tract
Velum-[–stop]	Nasal	Approximant for the nasal tract

In (6), I summarize how common manner features are interpreted in the present system. The feature [sonorant] (which is the same as non-obstruent) is not included, because it can be defined as [–stop] for Velum (i.e. a nasal) or [–stop, –fricative] under another articulator (i.e. an approximant).

(6) Interpreting traditional manner features

Traditional	Velum	Another articulator
Stop (plosive)	[+stop]	[+stop, –fricative]
Fricative	[+stop]	[–stop, +fricative]
Affricate	[+stop]	[+stop, +fricative]
Nasal (stop)	[–stop]	[+stop, –fricative]
Approximant	[+stop]	[–stop, –fricative]
Obstruent	[+stop]	[+stop] or [+fricative]

I have not included trills and flaps among manner features, because their contrasts with other manners remain to be worked out. For example, no language in P-base has a contrast between the trill [ʀ] and the fricative [ʁ]. If trills or flaps are confirmed to be a distinct manner category, another manner feature may need to be added.

6.2.3 Interpreting place features

Traditional place features refer to locations along the vocal tract, towards which articulators move. Some articulators can make forward or backward movement along the direction of the vocal tract, and such movement is represented with the [location] feature. If an articulator only moves perpendicular to the direction of the vocal tract, such as the velum or the tongue root, no [location] feature is given. On the basis of our examination of UPSID and P-base, three articulators can make forward or backward movement along the direction of the vocal tract, each having two or three contrastive locations. This is shown in (7).

(7) [location] values of three articulators

Articulator	[location] values	Traditional features
Lips (lower lip)	Front Central Back	Protruded Bilabial Labiodental
Tip	Front Central Back	Interdental Alveolar Retroflex
Body	Front Back	Palatal Velar

In previous studies, different terms are created for movements of similar direction. For example, the forward movements of Coronal, Dorsal, and Tongue-root are

[+anterior], [−back], and [+advanced] respectively (Halle 2003). The reasons seem to be that (i) the movements involve different articulators and (ii) there is no evidence that these movements yield a common acoustic or phonological effect. However, if articulators are specified, there is no need to use different terms for location features.

In the present analysis, the tongue root, the velum, and the glottis do not have a [location] feature, since they do not move along the direction of the vocal tract, but in the direction perpendicular to it (the Larynx movement will be discussed in Chapter 7). Nevertheless, their movements represent additional place features. In addition, the traditional place feature "palatal" can be viewed as involving two articulators at the same time, the tongue tip and the tongue body (Keating 1988; Browman and Goldstein 1989; Hayes 2009), and "uvular" can be viewed as involving the tongue body and the tongue root at the same time. Therefore, a more complete interpretation of traditional place features is shown in (8). For ease of exposition, I use two binary features [front] and [back] to represent location values. For each place feature, I have given a voiceless fricative as an example, as far as possible. A contrast between "protruded" bilabial and a regular bilabial is rare, but it has been proposed for two kinds of rounding in vowels, such as [y] vs. [ʉ] in Swedish (Fant 1983: 4; Ladefoged and Maddieson 1996: 269), although the two vowels may differ in "apicalization" as well (Fant 1983: 4).

(8) Interpreting place features (all articulators)

Articulator	Location values	Traditional places	Example
Lip (lower)	[+front, −back] [−front, −back] [−front, +back]	Protruded Bilabial Labiodental	? [ɸ] [f]
Tip	[+front, −back] [−front, −back] [−front, +back]	Interdental Alveolar Retroflex	[θ] [s] [ʂ]
Tip + Body		Palatal	[ɕ]
Body	[−back] [+back]	Palatal Velar	[ç] [x]
Body + Root		Uvular	[χ]
Root	[−advanced]	Pharyngeal	[ħ]
Glottis		Glottal	[h]

It is worth noting that although [front] and [back] are used for different articulators, there is no claim that there is something in common among the gestures. For example, there is no claim that [+front] of the lips (protruded) is related to [+front] of the tongue tip (interdental). Similarly, there is no claim that [+back] of the lower lip (labiodental) is related to [+back] of the tongue tip (retroflex) or the tongue body (velar).

It is widely assumed though that the manner feature [stop] has an intrinsic property regardless of what the articulator is. Similarly, it is widely assumed that the manner feature [fricative] has an intrinsic property regardless of what the articulator is. Such assumptions are supported by patterns of sound classes, such as the aspiration of voiceless stops in English.

A question then arises: Why is there a relation among manner gestures (features) made by different articulators, but a lack of such a relation among location gestures made by different articulators? A possible answer is that some sound classes are based on acoustic or perceptual properties (Steriade 1999; Hayes 1999). In particular, stops share a common acoustic property (a sudden change in speech signal or a period of silence), and it is possible that the acoustic property has an effect on sound classes. Similarly, fricatives share a common acoustic property (the presence of noise), which could have an effect on sound class, too. On the other hand, there is no known acoustic similarity among location gestures of different articulators, such as [+back] of the lower lip (labiodental) and [+back] of the tongue tip (retroflex), which may explain the lack of relation among such gestures.

6.2.4 *Other features*

There remain a few other features. [round] and [high] are traditional features, which can be seen as gestures, too. [lateral] is traditionally grouped with manner features; it is a gesture, too—the tongue tip is narrowed (and raised), which allows air to go through the sides. [stiff] is the gestural interpretation of voicing in consonants and tonal register in vowels (Halle and Stevens 1971), and [spread] is the gestural interpretation of [aspirated] or [breathy].

6.3 Representing vowels

Let us now consider the representation of vowels. As discussed in Chapter 4, there are sixteen basic vowels, which can be represented with four features. The eight front vowels are shown in (9). I have assumed some degree of underspecification (Steriade 1987; Archangeli 1988; 2011; Keating 1988; Hall 2007; Dresher 2009) and have only indicated features for the lips, the tongue body, and the tongue root.

(9) Representing basic vowels (back vowels omitted)

Articulators	Features	[i]	[y]	[ɪ]	[ʏ]	[e]	[ø]	[ɛ]	[œ]
Lips	[round]	–	+	–	+	–	+	–	+
Body	[high] [back]	+ –	+ –	+ –	+ –	– –	– –	– –	– –
Root	[advanced]	+	+	–	–	+	+	–	–

Next we consider non-basic vowels, which involve one or more of the additional features "nasal," "creaky" (or "glottal"), "breathy," and "retroflex." In (10) I illustrate how each additional feature is superimposed on [o], again with some degree of underspecification.

(10) Representing basic and non-basic vowels, exemplified with [o]

Articulators	Features	Basic [o]	Nasal [õ]	Creaky [oˀ]	Breathy [o̤]	Retroflex [o˞]
Lips	[round]	+	+	+	+	+
Tip	[back]					+
Body	[high] [back]	– –	– –	– –	– –	– –
Root	[advanced]	+	+	+	+	+
Velum	[stop]		–			
Glottis	[spread] [stiff]			 +	+ 	

Voiceless vowels (as used in whispered speech) can be represented too, but since they do not contrast with other vowels, I omit their representations.

6.4 Representing basic consonants

Consonants are traditionally described in terms of place of articulation and manner of articulation (plus voicing and aspiration features). We have discussed most manner features above, except "trill" and "flap." It is not clear how they are to be interpreted as articulatory gestures, and some feature studies have remained silent about them (e.g. Sagey 1986; Halle 1995; 2003). Harris (2006) suggests that the difference among regular, trill, and flap consonants does not lie in features, but in the prosodic environment. I shall, however, leave the topic open for now.

We have also discussed the representation of most place features. Let us take a close look at palatal consonants. P-base distinguished three palatal places, alveolar-palatal, pre-palatal, and palatal. In the present analysis, they involve two articulators each, as shown in (11).

(11) Analysis of alveolar-palatal (A-P), pre-palatal (P-P), and palatal (P)

Articulators	Features	A-P [ʃ]	P-P [ɕ]	P [ç]
Tip	[back] [fricative]	+ +	– +	– –
Body	[fricative]	+	+	+

First, [back] of the tongue tip can distinguish [ʃ] vs. [ɕ ç], where the tongue tip is raised in [ʃ] but not in [ɕ ç]. Second, [ɕ] is [+fricative] for both the tongue tip and the tongue body, whereas [ç] is [+fricative] for the tongue body only. In this representation, the tongue-body part of [ç] is a fricative and its tongue-tip part is an approximant, which means that [ç] is the same as [x^j]. I shall return to sounds with two articulators in Chapter 7.

6.5 Representing tones

Tones are contrastive features over vowels and, sometimes, over consonants. It is often thought that tone is the pitch pattern of the vowel (Fudge 1967: 4) or the voiced part of the syllable (Wang 1967; Chao 1968; Howie 1976; Hombert et al. 1979), but it can be argued that tone resides over the rime portion of the syllable only (Kratochvil 1970; Hyman 1985; Duanmu 1996). In addition, it is clear that in many cases each sound in the rime is a separate tone bearing unit, so that a long rime has two tone-bearing units (Woo 1969; Hyman 1985; Duanmu 1996). Consider the examples in (12) from Standard Chinese, where [á] and [ń] have a high tone (H) and [à] and [ǹ] have a low tone (L). In addition, a LH sequence is the same as a rise and a HL sequence is the same as a fall.

(12) Contrastive tones in Standard Chinese

Transcription	*Tone*	*Gloss*
[fáń]	HH (high)	‘turn (upside down)’
[fàń]	LH (rise)	‘sail’
[fàǹ]	LL (low)	‘opposite’
[fáǹ]	HL (fall)	‘rice’

In the representation, ‘turn’ and ‘sail’ minimally contrast in [á] and [à] and ‘turn’ and ‘rice’ minimally contrast in [ń] and [ǹ]. Therefore, it is reasonable to say that [á] and [à] are distinct phonemes, and so are [ń] and [ǹ] (Fu 1956). If this is the case, the number of vowel phonemes is (at least) doubled, and so is the number of tone-bearing consonants. On the other hand, if we value phonemic economy, according to which the number of phonemes should be minimized, the alternative is to list tones separately. In the above case, there are three phonemes [f], [a], and [n], plus four tones, HH, LH, LL, and HL—or, in an even simpler analysis, just two tones, H and L.

Although neither UPSID nor P-base includes tonal inventories, a feature theory ought to offer a representation of tones as well. Following previous studies, I assume that tones involve two articulatory dimensions, commonly called “register” and “pitch” (Yip 1980; 2002; Bao 1990a; 1999; Duanmu 1996; Kingston 2011). It has been proposed that “register” correlates with the stiffness of the vocal folds, and "pitch" correlates with their thickness (Halle and Stevens 1971; Zemlin 1998). If each feature has a two-way contrast, there are up to four level tones, shown in (13), where [stiff] represents the register feature (Halle and Stevens 1971) and [H] represents the pitch feature. Bao uses [H] and [L] for register values and [h] and [l] for pitch values.

(13) Feature representation of four level tones (A–D)

	A	B	C	D
Glottis-[stiff] (register)	+	+	–	–
Glottis-[H] (pitch)	+	–	+	–
Bao's notation	[H, h]	[H, l]	[L, h]	[L, l]

The feature for "register" is the same as that for voicing, following the interpretation of Halle and Stevens (1971). In addition, there is a well-known correlation between tonal registers and the voicing of consonants in the syllable onset during register genesis (the creation of tonal registers). The correlation is known as "voiceless high and voiced low" (Kingston 2011), whereby voiceless onsets lead to higher tones (upper register) and voiced onsets lead to lower tones (lower register). Besides the difference in tone levels, the upper register is realized with normal voice quality, and the lower register is realized with a murmured voice quality.

The feature for "pitch" is related to the thickness of the vocal folds, where [+H] (high) is produced when the vocal folds are stretched and thin, and [–H] (low) is produced when the vocal folds are thick and not stretched. It has been proposed that the larynx plays a role, too, where it is raised in [+H] and lowered in [–H] (Hombert 1978).

Bradshaw (1999) proposes a slightly different model. She argues that [voiced] and low tone involve the same feature. The difference lies in what node the feature is linked to. If the feature is linked to the Laryngeal node, it is similar to [voice] in a consonant (or a low register in a vowel). If it is linked to a mora, it is similar to a low tone. Presumably, if it is linked to both the Laryngeal node and a mora, it is similar to a low tone of a low register. It can be seen that Bradshaw's model does not assume a gestural interpretation of features, and is less constrained by predicting far more possible feature structures.

6.6 Feature specification (underspecification)

It is clear that not every feature is specified for every sound. For example, [m] can be referred to as "bilabial nasal stop." A theoretical question then is: how many features need to be specified for a given sound? Consider the feature specification for [m] in English, shown in (14).

(14) Feature specification for [m] in English

	Feature	Traditional term
Required	Lips-[+stop]	Labial stop
	Velum-[−stop]	Nasal
Predictable	Glottis-[−stiff]	Voiced
	Glottis-[−spread]	Unaspirated
Unclear	Body-[?high]	High or not?
	Root-[?advanced]	Pharyngealized?
	Glottis-[?H]	H tone or not?

As far as English is concerned, Lips-[+stop] is required to distinguish [m] from [n], the latter being Tip-[+stop]. In addition, Velum-[−stop] is required to distinguish [m] from [b], the latter being Velum-[+stop]. The next two features, "voiced" and "unaspirated," are predictable, because English nasals are always voiced (except in whispered speech) and unaspirated. Beyond these, it is hard to say what values other features should be. For example, what is the height of the tongue body? Is [m] pharyngealized or not? What is its tone value?

In practice, few phonologists fully specify every feature for every sound. For example, in the IPA system, consonants are specified for consonant features only, and vowels are specified for vowel features only.

A sound is "underspecified" if it is only specified for some features but not others, and the theory to determine how this is done is the theory of underspecification (Halle 1959; Stanley 1967; Steriade 1987; 1995; Archangeli 1988; 2011; Keating 1988; Inkelas 1995; Rice 1995; Dresher 2009).

Steriade (1987) points out that a specified feature can be a source of feature spreading (assimilation), while a sound unspecified for the feature tends either to be a target of feature spreading or to let the feature pass through it (i.e. being "transparent" to feature spreading). In addition, there is a general consensus that specified features resist change, while unspecified features are a source of variability. For example, if a vowel is specified for a high tone, it tends to be realized as such, but if a vowel is unspecified for a tone, it tends to have variable pitch levels depending on the environment (Liberman and Pierrehumbert 1984; Beckman and Pierrehumbert 1986). Similarly, if a sound is specified for the height and backness of the tongue, it tends to be realized as such, but if a sound is unspecified for these features, it tends to have variable tongue positions (Keating 1988; Inkelas 1995; Rice 1995).

All linguists agree that contrastive sounds must be distinguished. There is some disagreement on how this is achieved. Let us consider three approaches, whose procedures are interpreted in (15)–(17).

(15) Contrastive Specification (Steriade 1987)
   a. Fully specify all sounds in an inventory (for a given set of features).
   b. If two sounds differ in only one feature, specify this feature for both of them.
   c. Leave all other feature values unspecified.

(16) Radical Underspecification (Archangeli 1988)
   a. Choose any feature and specify one value for the sounds of an inventory.
   b. Repeat the process till every sound has an unique set of values.

(17) Successive Division Algorithm (Dresher 2009)
   a. Choose any feature whose values can divide an inventory into two subsets.
   b. Choose another feature whose values can divide a subset into two.
   c. Repeat till every subset has only one member.

Let us consider how each approach would specify a simple vowel inventory [i a u]. Three scenarios of Contrastive Specification are shown in (18)–(20).

(18) Contrastive Specification of [i a u], using [high], [round], and [nasal]

	Full specification			*Underspecification*		
	[i	a	u]	[i	a	u]
[high]	+	−	+	+	−	
[round]	−	−	+	−		+
[nasal]	−	−	−			

(19) Contrastive Specification of [i a u], using [high] and [back]

	Full specification			*Underspecification*		
	[i	a	u]	[i	a	u]
[high]	+	−	+		−	+
[back]	−	+	+	−		+

(20) Contrastive Specification of [i a u], using [high], [round], and [back]

	Full specification			*Underspecification*		
	[i	a	u]	[i	a	u]
[high]	+	−	+			
[round]	−	−	+			
[back]	−	+	+			

In (18), three features are used, where [high] is contrastive for [i a], [round] is contrastive for [i u], while [nasal] is not contrastive for any vowel. The solution shows that every vowel has a unique set of specified feature values. In (19), two features are used, where [high] is contrastive for [a u] and [back] is contrastive for [i u]. The solution, which is different from that of (18), shows again that every vowel has a unique set of specified feature values. In (20), three features are used, but every pair of sounds differ in two features and no feature serves as the sole contrast. Therefore, there is no solution. The examples show that Contrastive Specification can have alternative solutions when different sets of features are used, but if too many features are used, there might be no solution.

Next we consider Radical Underspecification of the inventory [i a u], using the features [high] and [round] only, shown in (21).

(21) Radical Underspecification of [i a u], using [high] and [round]

	Solution A			*Solution B*			*Solution C*			*Solution D*		
	[i	a	u]	[i	a	u]	[i	a	u]	[i	a	u]
[high]		–		+		+		–		+		+
[round]	–	–		–	–				+			+

If we use two features, there are four solutions, since we can use either the plus or the minus value of each feature. In some solutions, one of the sounds can be completely unspecified, such as [u] in solution A and [i] in solution C. This is supposed to be the default sound, the one most likely to be added when an extra sound is needed (Archangeli 1988). Three differences can be observed between Contrastive Specification and Radical Underspecification. First, in general, Contrastive Specification uses more specified feature values than Radical Underspecification. Second, Contrastive Specification has fewer solutions than Radical Underspecification. Specifically, for a given set of features, Contrastive Specification has just one solution, while Radical Underspecification has up to 2^n solutions, where *n* is the number of relevant features. Third, Contrastive Specification may have no solutions when too many features are involved, whereas Radical Underspecification has no such problem, because when a solution is reached, i.e. when every sound has a different column of feature values, the process stops and additional features are not used.

Finally let us consider the analysis of the inventory [i a u] by the Successive Division Algorithm, using [high] and [round], shown in (22).

(22) Analysis of [i a u] by the Successive Division Algorithm

	Solution A				*Solution B*		
	[i	u	a]		[i	a	u]
[high]	[+	+]	[–]	[round]	[–	–]	[+]
[round]	[–]	[+]		[high]	[+]	[–]	

It can be seen that the same two features can yield two different solutions. In other words, the solution is sensitive to the order (or "hierarchy") in which the features are used (which is not the case in Radical Underspecification). In general, given *n* relevant features, there are *n* factorial solutions in Successive Division Algorithm, which exceeds those in Radical Underspecification when there are four or more features.

It can be seen that, in all three approaches, feature specification is inventory-driven, in that the larger the inventory, the larger the average number of specified feature values per sound. In addition, in every approach, alternative solutions are possible for a given inventory, although Contrastive Specification allows fewer alternative solutions than the other two. Finally, Radical Underspecification offers fewer specified feature values than the other two.

It is hard to evaluate the predictions of various approaches to underspecification. For example, the Successive Division Algorithm offers many different ways to specify the "same" inventory of sounds, such as the vowel inventory [i e a o u]; accordingly, it predicts many different patterns in which these vowels may behave in different languages, such as whether a vowel will trigger feature spreading, and how much variability there is in the realization of each vowel. But relevant data to evaluate the predictions are hard to come by. Besides, there are other views of underspecification. For example, Inkelas (1995) and Ito et al. (1995) argue that feature specification is not inventory-driven. In addition, they argue that there is no need to minimize specified feature values. Despite the disagreement, all linguists agree that contrastive sounds must be distinguished, that some features are unspecified, and that different languages can use different specifications for what appears to be the same phoneme inventory.

6.7 Phonetic variation of features

A theory of features needs to account for two well-known facts. First, the number of contrasts (or the features needed to represent them) is rather small. Second, the number of non-contrastive phonetic variations of a sound is rather large and often language-specific.

For example, we have seen in UPSID and P-base that there are only sixteen basic vowels, representable with four binary features (Chapter 5). We have also seen that the phonetic realization of a sound can vary considerably. For example, consider Fig. 2.3 again, repeated here as Fig. 6.1, which shows the phonetic variation of [i a u] in context, where context may include stress, prosodic position, and influences from adjacent sounds. The height and backness of the tongue are reflected by F1 and F2 measurements.

Keating (1985) suggests that there are two approaches. One is to enrich the feature system so that it can represent both contrastive and non-contrastive differences. The other is to keep a simple feature system for contrastive differences only, supplemented by a separate (language-specific) phonetic system.

Ladefoged (1972) takes the first approach. He argues that features should represent not only contrastive differences but non-contrastive differences as well. The solution is to see features not as binary contrasts but as phonetic scales, each of which can be divided in different ways. For example, a language can choose how many degrees of contrast the scale [back] is divided into, and where the dividing lines are (or where the center of each contrastive category is). A similar view is expressed by Kohler (1966: 339), who proposes a number of scales. Some scales are discrete and have two or three values each, such as airstream mechanism, which has three values "pulmonic-velaric-glottalic." Other scales are continuous ("place," "stricture,"

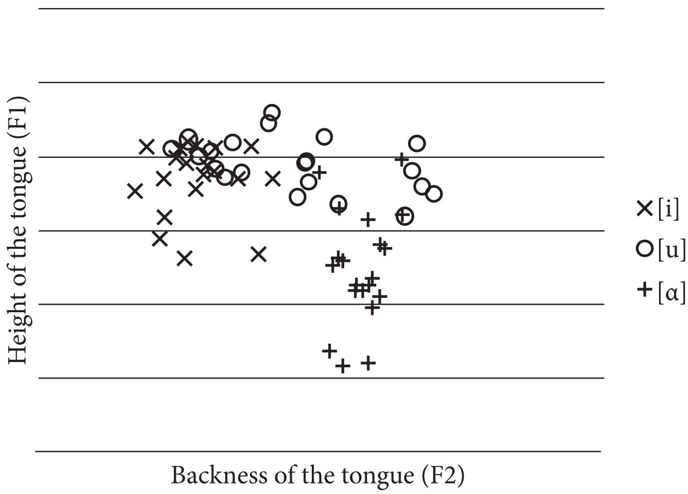

FIG. 6.1 Contextual variation of three vowels (twenty tokens each) by one female speaker of American English. The vowels were narrowly transcribed as [i], [ɑ], or [u], spoken by speaker s0101a from Columbus, Ohio, in the Buckeye Corpus (Pitt et al. 2007; measured by San Duanmu).

"time," "pitch," and "loudness"), each having "an infinite number of sub-divisions." The challenge for this approach is to explain why there are so few contrasts. For example, why are there just sixteen basic vowels? Why is there no more than a binary contrast for most features (if not all)?

Flemming (1995) also takes the first approach. Like Ladefoged, Flemming believes that features are phonetic scales and that features should represent both contrastive and non-contrastive differences. Flemming further proposes that (i) a feature scale can be divided into many degrees and (ii) each language can impose a requirement on the minimal distance between two contrastive categories, measured by the number of degrees between them. For example, if we divide the feature [back] into ten degrees, and require that there be a minimal distance of four degrees between two contrastive locations, then there are at most three contrastive categories in backness, but four ways to place them. This is shown in (23).

(23) Four possible systems (L1–L4) of three-way contrast (A–C) in the backness scale, if the scale has ten degrees of just noticeable differences (1–10) and a minimal distance of four degrees is required between two contrastive locations (Flemming 1995)

	1	2	3	4	5	6	7	8	9	10
L1	A	.	.	.	B	.	.	.	C	.
L2	A	.	.	.	B	.	.	.	.	C
L3	A	.	.	.	.	B	.	.	.	C
L4	.	A	.	.	.	B	.	.	.	C

If we change the minimal distance requirement from four to three degrees, then more systems of three-way contrast are possible. These are shown in (24).

(24) Possible systems of three-way contrast (A–C) in the backness scale, if the scale has ten degrees of just noticeable differences (1–10) and a minimal distance of three degrees is required between two contrastive locations (Flemming 1995)

	1	2	3	4	5	6	7	8	9	10
L1	A	.	.	B	.	.	C	.	.	.
L2	A	.	.	B	.	.	.	C	.	.
L3	A	.	.	B	.	.	.	.	C	.
L4	A	.	.	B	.	.	.	.	.	C
L5	A	.	.	.	B	.	.	C	.	.
L6	A	.	.	.	B	.	.	.	C	.
L7	A	.	.	.	B	.	.	.	.	C
L8	A	.	.	.	.	B	.	.	C	.
L9	A	.	.	.	.	B	.	.	.	C
Etc.										

The notion of a minimal distance explains why there are fewer contrastive categories than non-contrastive phonetic differences, and why different patterns are possible for the same number of contrasts. On the other hand, Flemming fails to explain why there are so few contrasts, or to offer a prediction of the maximal number of possible contrasts. In fact, Flemming seems to assume that the minimal distance between two contrastive categories could be as small as a language prefers. If so, the predicted number of possible contrasts would far exceed what we have evidence for.

Chomsky and Halle (1968) take the second approach, in which features represent contrastive differences only. They also claim that the phonetic realizations of features are predictable from "universal rules." However, few details of such rules are offered. In fact, phonetic studies have shown that phonetic realizations of features are not universal but often language-specific (e.g. Ladefoged 1972; Disner 1983; Keating 1985). For example, consider Fig. 1.3 again, repeated as Fig. 6.2, which shows small but systematic differences between [i y e ø] in Norwegian and their counterparts in German. Such language-specific differences are not accounted for by Chomsky and Halle (1968).

Keating (1985) also follows the second approach. Unlike Chomsky and Halle (1968), Keating recognizes language-specific variations, such as those in Fig. 6.2. In addition, Keating argues that even contextual variations, such as those in Fig. 6.1, can be language-specific. She proposes, therefore, that we need not only a simple feature system than represents phonological contrasts but also a language-specific phonetic system that describes how features are realized (see also Beddor et al. 2002).

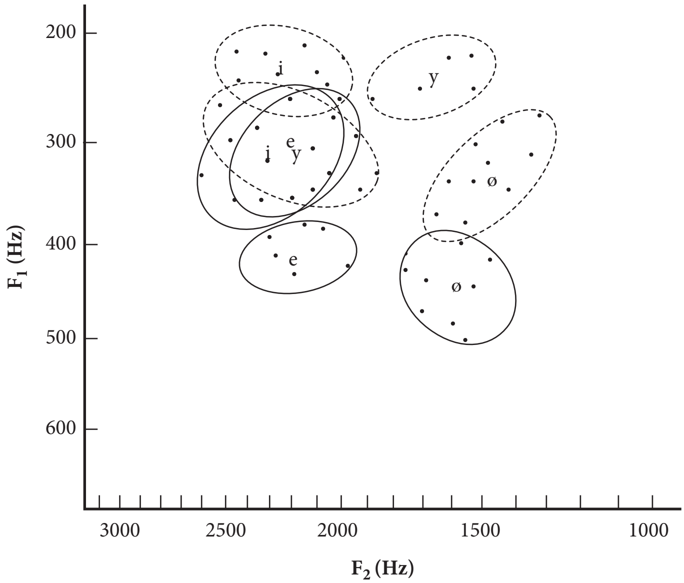

FIG. 6.2 Cross-language differences between [i y e ø] in Norwegian (solid line, 10 speakers) and their counter-parts in German (dotted line, 6 speakers), from Disner (1983: 67).

Keating's proposal seems to have addressed both problems. The phonological component of features is simple, probably universal, and accounts for the small number of contrasts. The phonetic component is complex, language-specific, and accounts for the large number of phonetic variations. However, some important questions remain, shown in (25).

(25) Questions for feature theory
  a. Why is the contrastive feature system so simple?
  b. Why is there so much phonetic variation in the realization of features?
  c. Why is the phonetic variation of features language-specific?

Let us now consider how the present proposal answers these questions. With regard to (25a), we have seen in section 6.1 that the number of features, or phonetic scales, is limited. In addition, a binary contrast offers a more robust distinction than three or more degrees of contrast. Specifically, a two-way contrast requires judging the direction of change only, whereas a three-way contrast requires judging the degree of change as well. For example, consider a binary contrast in tones (low vs. high). If the tone of a given syllable is low, we can determine the next tone by checking whether the pitch (or Fo) is going up. Now suppose we have a three-way contrast in tones (low, mid, and high). If the current tone is low and the pitch of the next tone is going up, we still do not know whether the next tone is mid or high, because we still need to determine how much the pitch has gone up, and this is quite difficult, because there is considerable variation in pitch range among speakers and in context.

Next we consider why there is so much phonetic variation in the realization of features. There seem to be two reasons. First, binary contrasts leave much room for variation. For example, to change from a low tone to a high tone, all one needs to do is to raise the pitch, whether it is just a few hertz, or dozens of hertz. Second, many factors can affect the phonetic result, such as coarticulation, prosody, the mental state of the speaker, and the energy level of the speaker. The fact that a gestural form can have variable physical realizations is not unique to articulation. Let us compare articulation with locomotion, shown in (26).

(26) Gestural form and its (external) physical realization

	Gestural form	Physical measure
Articulation	Body-[high]	Tongue height
Locomotion	Walking	Speed

As [high] of the tongue body is a distinct articulatory gesture, walking is a distinct gestural form of locomotion (in contrast to running or crawling). In addition, as the realization of [high] is contextually variable, the speed of walking is affected by many factors, such as the road condition and the physical state of the walker. It is possible, for example, that one's walking speed in some conditions is faster than one's running speed in other conditions, although under the same conditions walking is slower than running. In fact, it would be highly unusual if someone walks at exactly the same speed, regardless of external conditions or that individual's physical and mental state. Similarly, it would be highly unusual if features were always realized in exactly the same way, regardless of contextual factors.

Finally, let us consider why phonetic variations of features are often language specific. There are three possible reasons. First, as discussed in the preceding section, what appears to be the same vowel (or the same phonetic transcription) may in fact have different feature specifications in different languages; therefore, we expect them to be realized in different ways. Second, the physical and cultural environment of a speech community is likely to play a role. For example, in studies on "the pace of life", it has been found that people in different cities walk at different speeds (Lowin et al. 1971; Bornstein and Bornstein 1976; Levine and Norenzayan 1999), and a number of factors have been suggested: population size, culture, productivity, and climate. If so, similar factors, or additional ones, may very well affect how people talk. Third, several studies have found that people in a speech community imitate each other in various phonetic details (Sancier and Fowler 1997; Goldinger 1998; Nielsen 2008; 2011). Given such effects (physical and cultural environment and phonetic imitation), and the fact that binary contrasts leave much room for variation, it would be highly unusual if language-specific patterns did not emerge. Similarly, it would be highly unusual if people in all places walked at exactly the same speed, or if every individual kept his or her speed constant regardless of his or her environment.

6.8 Summary

I have reviewed several feature theories and offered a minimally sufficient feature system in which features are gestures of active articulators. The new system assumes just seven articulators (including Larynx, to be discussed in Chapter 7)—the minimum any theory must assume, and far fewer features than previously proposed e.g. by Halle (2003) and Ladefoged (2007). I have also shown that the system is sufficient in that it can distinguish all contrasts in UPSID and P-base.

I have shown, too, that the proposed system can help answer some theoretical questions—in particular why there are so few contrasts, why there is so much phonetic variation in the realization of features, and why phonetic variation of features is often language-specific.

Since the present feature system is based on contrast, we might wonder whether more features are needed once we consider sound classes as well. Let us consider nine traditional features not used in the present system: [continuant], [constricted glottis], [sonorant], [consonantal], [constricted pharynx], [strident], [low], [distributed], and [slack vocal folds]. Let us consider whether this has to be the case. The first three are replaced with different names: [continuant] is [stop], [+sonorant] is [–stop, –fricative] (or Velum-[–stop]), and [+constricted glottis] is Glottis-[+stop]. The feature [+constricted pharynx] is likely the same as [–ATR]. The feature [–consonantal] refers to "dorsal articulation" (Chomsky and Halle 1968: 302), which means using Body but without [+stop] or [+fricative] values under it. The feature [slack vocal folds] is originally used with [stiff vocal folds] to create a three-way contrast in voicing or glottal tension (Halle and Stevens 1971), but such a contrast is not found in the present study. The feature [strident] has been used for two purposes. One is to represent affricates, which in the present study are [+stop, +fricative]. The second purpose is to distinguish alveolar, retroflex, and palatal fricatives from interdental and labial fricatives (and possibly from velar and uvular fricatives too). In the present analysis, "strident fricatives" [s ʂ ʃ] are Tip-[–front, +fricative], and other fricatives are non-strident; this accounts for the fact that the English plural suffix shows up as [ɪz] after [s ʃ] but as [s] after other voiceless fricatives (English lacks [ʂ]). The feature [distributed] is originally used for palatals; in the present analysis, it is replaced by the co-presence of Tip and Body. Finally, I have argued that [low] is not needed, either for contrast or for sound classes. Thus, it remains to be seen what additional features are needed when we consider sound classes.

7

Complex sounds

I use "complex sound" to refer to anything that is thought to be a single sound but seems more complicated than a regular sound. It includes transcriptions with two or more IPA symbols, such as [ts], [ks], [nd], [kw], [pr], [k͡p], and [õ̤]. It also includes transcriptions with a single IPA symbol, such as [!] and [ɓ], which seem to require sequential gestures.

My term covers what Sagey (1986) calls "contour segments," which contain opposite values of the same feature (called a "contour feature"), such as [ts], which contains [−continuant, +continuant] and [nd], which contains [+nasal, −nasal]. My term also covers what Sagey (1986) calls "complex sounds," which do not involve contour features but "simultaneous multiple articulations," such as [k͡p]. Moreover, my term covers sounds that have simple transcriptions but seem to involve contour features, such as [!] and [ɓ]. I use a more general term, rather than Sagey's terms, because it is not always obvious which sounds do or do not require contour features. For example, according to Sagey (1986), clicks do not involve contour features, but some linguists believe they do. On the other hand, Sagey considers affricates to involve contour features, but some linguists believe they do not.

Complex sounds call for special attention because they raise a theoretical question: How do we define what can be a single sound and what cannot? Much inconsistency can be seen in the literature. For example, [ts] is treated as two sounds in English, while [tʃ] is treated as one (Ladefoged and Johnson 2011). Similarly, Wiese (1996) treats [ts] and [tʃ] as single sounds in German, while Moulton (1956) and Kohler (1999) treat them as two sounds each. Moreover, according to Sagey (1986) and Halle (1995; 2003), homorganic affricates are possible, such as [pf], [ts] and [tʃ], but non-homorganic affricates are not, such as [ps] and [ks]; however, the assumption is not shared by others (e.g. Wiese 1996), and [ps] is found as a phoneme in some languages in P-base.

I shall propose that the answer to the theoretical question lies in feature compatibility, according to which some complex sounds are single sounds and some not. I start with a discussion of feature compatibility, followed by complex sounds whose features are compatible, which includes affricates, consonant–glide units, and consonant–liquid units. Next I discuss complex sounds whose features are

A Theory of Phonological Features. First edition. San Duanmu.

incompatible, which include contour tones and pre- and post-nasalized stops. Finally, I discuss non-pulmonic complex sounds, whose representations have been problematic. I shall show that they can be represented as single sounds in some contexts, such as when they occur between vowels.

7.1 Compatibility of feature values: The No Contour Principle

In Chapter 1, I defined sounds in terms of time and feature values. The definition is repeated here in (1).

(1) Sounds defined by time
A sound is a set of compatible feature values in one time unit.

Feature values are compatible if they do not involve "contour features." This constraint can be specified in (2)–(4), after Duanmu (1994).

(2) No Contour Principle
A sound cannot contain contour feature values (or sequential feature values).

(3) Contour feature values
A and B form contour feature values if (i) A and B are opposite values of the same feature and (ii) A and B are performed by the same articulator in the same time unit.

(4) Representation of contour feature values ([F] is a feature)

```
Time unit                      X          or       X
                               |                   |
                          Articulator         Articulator
                              /  \                / \
Contour feature values →   [+F]  [–F]          [–F] [+F]
```

The No Contour Principle is implicitly assumed in Chomsky and Halle (1968), and Hoard (1971) calls it the "principle of simultaneity," namely, all feature values within a sound must be simultaneously executable and not sequentially ordered. In a number of cases, however, it was thought that contour features cannot be avoided.

Contour features have been used to represent contour tones, where a rise is [–H, +H] and a fall is [+H, –H] (Goldsmith 1976; Williams 1976). They have also been used to represent other things, such as pre- and post-nasalized stops (e.g. Anderson 1976; Sagey 1986; Kehrein 2006) and affricates (e.g. Hoard 1971; Campbell 1974; Sagey 1986; Steriade 1993). However, a main problem with contour features is that it leads to excessive over-prediction of possible sounds (Duanmu 1994). In addition, most sounds do not involve contour features. Moreover, for those that do, not enough attention has been paid to their other properties, such as whether a vowel that carries a contour feature is a single sound or a two-sound cluster. For example, Duanmu (1994) argues that, as far as Chinese languages are concerned, all contour tones fall on long rimes (those that have two time units each), where each time unit carries just one tone value (a level tone), in agreement with an earlier proposal of Woo (1969). This is illustrated in (5) with a rising tone.

(5) Rising tone and vowel length

Rise on a short vowel	Rise on a long vowel
[a] X | Glottis /\\ [–H] [+H]	[a:] X X | | Glottis Glottis | | [–H][+H]

When a rise falls on a short vowel, it is a contour feature. When a rise falls on a long vowel, which has two time units, no contour feature is needed. Thus, a crucial point of interest is whether a contour unit is short or long. There is some evidence for the No Contour Principle. For example, in a perception study Greenberg and Zee (1979) have found a strong positive correlation between the length of a vowel and the perceived degree of "contouricity" of its tone, where "contouricity" refers to the steepness of rise in pitch. For example, a 50Hz rise on a short vowel (90ms) was heard as nearly a level tone, whereas a 50Hz rise on a long vowel (250ms) was heard as a sharp rise, even though the former is visually a sharper rise. Such results raise serious questions as to whether contour features, which ought to fall on short sounds, are physically possible. As the null hypothesis, I shall assume that they are not.

The No Contour Principle is similar to the No Branching Constraint of Clements and Hume (1995: 255). Their definition is shown in (6), slightly rephrased to match the current terminology.

(6) The No Branching Constraint (Clements and Hume 1995: 255)
Configurations of the form below are ill-formed, where A is any class node (including the root node), A immediately dominates B and C, and B and C are identical articulators, or values of the same feature.

General form	Examples		
A /\\ B C	C-place /\\ Coronal Coronal	Coronal /\\ [+anterior][+anterior]	Coronal /\\ [–anterior] [+anterior]

However, Clements and Hume (1995) allow two (or more) "root" nodes to occur within a timing unit (i.e. within a sound). This is shown in (7), where X is a timing unit.

(7) Good and bad contour units for Clements and Hume (1995: 254)

Good	Bad
X /\\ Root Root | | [+nasal] [−nasal]	X | Root /\\ [+nasal] [−nasal]

Because a “good” contour unit takes only one timing unit, Clements and Hume (1995) predict that short contour tones are possible, which I have just argued not to be the case.

The No Contour Principle is a constraint on feature structure, not on IPA symbols. For example, consider the structures in (8), where [−stiff] is the gesture for “voiced” and the palatal [j] has both the articulator Tip and the articulator Body (Chapter 6).

(8) Feature representation of [g], [j], and [gj]

[g]	[j]	[gj]
	Tip-[−stop, −fricative]	Tip-[−stop, −fricative]
Body-[+stop, −fricative]	Body-[−stop, −fricative]	Body-[+stop, −fricative]
Glottis-[−stiff]	Glottis-[−stiff]	Glottis-[−stiff]

In terms of IPA symbols, where [gj] is made of [g] and [j], [gj] seems to contain a contour feature: Body-[+stop] of [g] and Body-[−stop] of [j] seem to form a contour feature [+stop, −stop]. In terms of feature structure, however, none of the structures violates the No Contour Principle. In addition, in [gj], the co-presence of Tip and Body implies a palatal component, and the presence of Tip-[−stop, −fricative] implies an approximant component. Therefore, the best interpretation of the structure for [gj] is a palatalized [g], which is what [gj] is meant to refer to.

Needless to say, whether two sounds will merge into a complex sound or not depends on the context, even if they have compatible features. For example, [k] and [p] can form a complex sound [k͡p], shown in (9), where [+stiff] is the gesture for “voiceless.”

(9) Feature representation of [k], [p], and [k͡p]

[k]	[p]	[k͡p]
	Lips-[+stop, −fricative]	Lips-[+stop, −fricative]
Body-[+stop, −fricative]		Body-[+stop, −fricative]
Glottis-[+stiff]	Glottis-[+stiff]	Glottis-[+stiff]

The gestures for [k] and [p] are can be performed simultaneously, giving the complex sound [k͡p]. However, in English [k] and [p] do not merge into [k͡p] but stay as a sequence of two separate sounds. For example, in *napkin*, the gesture for [k] comes after that for [p]. Similarly, [s] and [w] can form a compatible complex sound [s^{w}] (or [$s^{\underset{\circ}{w}}$]), as shown in (10).

(10) Feature representation of [s^{w}]
Lips-[+round]
Tip-[–stop, +fricative]
Glottis-[+stiff]

In [s^{w}] the gesture for lip rounding occurs at the same time as the gestures for [s]. In contrast, in [sw] the gesture for lip rounding occurs after the gestures for [s]. There is indeed such a difference between Chinese and English. As Chao (1934: 373) observes, in the English word *sway*, [s] "is not at all labialized for most of its duration," whereas in the Chinese word [s^{w}ei] 'year', [s] "is completely labialized."

The No Contour Principle looks like a structural constraint, but it probably originates from a physical one, namely, an articulator cannot act fast enough to perform two opposite gestures in one unit of time (and the ear probably cannot process two opposite values of the same feature in one unit of time either). Indeed, there are other physical constraints. For example, the velum cannot perform [fricative] and the tongue root cannot perform [stop]. To the list we can add the lack of simultaneous labio-dental ([+back] for the lower lip) and interdental ([+front] for the tongue tip). Physically, it is possible to move the lower lip to a position between the teeth (labio-dental), while at the same time moving the tongue tip to the same position (interdental). But such a configuration does not offer enough acoustic difference from a simpler one: If the tongue tip is above the lower lip, the acoustic effect is similar to a simple interdental, and if the tongue tip is below the lower lip, the acoustic effect is similar to a simple labio-dental.

7.2 Complex sounds whose features are compatible

There are three cases in this category: affricates, consonant–glide units, and consonant–liquid units.

7.2.1 Affricates

Affricates have been a problem for feature theory, in that they seem to require a "contour feature," or sequential feature values within a sound. For example, consider the affricate [ts]. The [t] part is a stop, traditionally defined as a complete closure in the oral tract, whereas the [s] part is a fricative, traditionally defined as an incomplete closure in the oral tract. This leads to a contour feature [+stop, –stop], where [+stop] precedes [–stop], as shown in (11).

(11) Representation of [ts] with contour features (Halle and Clements 1983; Sagey 1986)

[t]	[s]	[ts]
[+stop]	[−stop]	[+stop, −stop]

Clements and Hume (1995) offer a slightly different representation, shown in (12), in which a stop and a fricative are sequentially ordered, but share one timing slot. A similar representation is offered by Selkirk (1982) and Wiese (1996).

(12) Representation of [ts] as two sounds sharing one position (Clements and Hume 1995)

```
 X
 /\
t  s
```

The reverse ordering is thought to be possible as well, i.e. [st] is predicted to be a possible single sound. It can be seen that (11) and (12) are essentially equivalent: They both require sub-segmental timing, or contrastive use of sequential ordering within a time unit.

Some linguists have attempted to avoid contour features. For example, in Jakobson et al. (1952), affricates are [+strident] stops, where [strident] refers to turbulence or noise in a fricative. The idea is adopted by many others, such as Steriade (1989), Clements (1999), Kehrein (2006), Hall (2007), Lin (2011). A shortcoming of this approach is that [strident] is rather ad hoc (Sagey 1986: 281), since it is not an articulatory gesture, whereas most other features are (or can be interpreted as such).

Chomsky and Halle (1968: 318) do not use contour features for affricates either. Instead, they use the feature [delayed release], so that [t] is [−delayed release] and [ts] is [+delayed release]. In [ts] the release of the stop is slow, and "turbulence is generated in the vocal tract so that the release phase of affricates is acoustically quite similar to the cognate fricative." In contrast, in [t] the release of the stop is sudden, which "is normally accompanied by much less or no turbulence." However, [delayed release] is a contour feature in disguise, because it calls for sequential articulatory gestures, and some linguists find it problematic (Anderson 1974; Campbell 1974; Sagey 1986). In addition, as Hoard (1971) points out, [delayed release] cannot represent the formation of affricates. For example, [t] and [s] can often form an affricate [ts]. However, both [t] and [s] are [−delayed release], yet [ts] is [+delayed release]. It is hard to explain how two negative values of a feature can merge into a positive value.

Hoard (1971), Halle and Clements (1983), and Sagey (1986) propose that affricates involve the contour feature [−continuant, +continuant]. A similar proposal is made by Steriade (1993), where affricates have two aperture values [A_o A_f], sequentially ordered, where A_0 is the aperture of a stop and A_f is the aperture of a fricative. Besides the problem of over-prediction, the representation has another problem: It predicts an "edge effect" (Sagey 1986), according to which affricates behave like stops to their left and like fricatives to their right. However, Hualde (1988) argues that the prediction is incorrect, because an affricate can behave like a stop or a fricative either to its left or to its right.

Hualde (1988), Lombardi (1990), and Padgett (1995) propose that affricates are [–continuant, +continuant], where the two feature values are not sequentially ordered at an abstract level, although they are ordered at the phonetic level. However, it is unclear how two contradictory feature values can be simultaneously present. In addition, it is unclear why most features do not allow contradictory values.

I would like to propose that affricates involve two independent gestures, [stop] and [fricative], rather than two conflicting feature values, [–continuant] and [+continuant]. Specifically, unlike traditional definitions of [fricative], which focus on the degree of closure, I suggest that the definition should focus on the manner of closure instead. It has been observed that the tongue has a "groove" shape in fricatives, with a lowered center and raised edges, based on 3-D imaging of the tongue surfaces and electropalatography (Stone and Lundberg 1996). Therefore, I propose that [fricative] refer to the edge closure of an articulator. In [ts], the edges of the tongue exert extra force during closure, which leads to a momentary delay in their release, giving the fricative effect. Because [stop] and [fricative] are independent gestures, they are free to combine, without creating a contour feature. The analysis of [t], [s], and [ts] are shown in (13).

(13) Feature representation of [t], [s], and [ts]

[t]	[s]	[ts]
Tip-[+stop]	Tip-[+fricative]	Tip-[+stop, +fricative]
Glottis-[+stiff]	Glottis-[+stiff]	Glottis-[+stiff]

Our analysis is free from the problems in previous ones. First, it does not need contour features. Second, it does not use ad hoc features, such as [strident] and [delayed release]. Third, our definition of [fricative] is a more accurate description of the gesture. Finally, our definition explains why [fricative] is available for the lips, the tongue tip, and the tongue body, but not for the velum, the tongue root, the glottis, or the larynx (Chapter 6): The former three articulators can make edge closures, but the latter four articulators cannot.

In most common affricates, the stop and the fricative share the same articulator, such as [ts] and [kx]. It is less clear whether a stop and a fricative with different articulators can form an affricate, such as [ps] and [ks]. In Sagey (1986) and Halle (2003), such affricates are impossible (because they assume that two articulators cannot have different manner features—an issue not relevant for us here). In the present analysis, however, affricates like [ps] and [ks] are possible, because they do not violate the No Contour Principle. The representation of [ks] is shown in (14). Similarly, [ps], [px], [tf], [tx], [kf], etc. are all possible complex sounds.

(14) Feature representation of [ks]
Tip-[+fricative]
Body-[+stop]
Glottis-[+stiff]

It is relevant to consider another kind of "affricate," which is made of a nasal and a fricative, such as [mz]. Consider the feature structure in (15).

(15) Feature representation of [m], [z], and [mz]

[m]	[z]	*[mz]
Lips-[+stop, –fricative]		Lips-[+stop, –fricative]
	Tip-[–stop, +fricative]	Tip-[–stop, +fricative]
Velum-[–stop]	Velum-[+stop]	Velum-[+stop, –stop]
Glottis-[–stiff]	Glottis-[–stiff]	Glottis-[–stiff]

It is reasonable to assume that fricatives ought to be non-nasal (Velum-[+stop]). The phonetic reason is that, in order to create the acoustic effect of a fricative, a certain amount of air pressure is required in the oral tract, and opening the nasal tract removes that air pressure (Ohala 1975; Ohala and Ohala 1993). If so, [mz] is not a possible sound, because Velum has sequential values for [stop], which create a contour feature. The conclusion agrees with the fact that such "affricates" are rare in UPSID or P-base and are open to alternative analyses.

7.2.2 CG (consonant–glide) units

A CG unit is made of a consonant and a glide, such as [k^w]. The consonant is called the "primary articulation" and the glide, often written in superscript, is called the "secondary articulation." A CG unit seems to contain sequential feature values. For example, in traditional features, [k^w] seems to call for the representation in (16).

(16) Literal representing of [k^w] in traditional features

	[k	^{w}]	
[stop]	+	–	← contour feature
[round]	–	+	← contour feature
[voice]	–	+	← contour feature

However, it is possible to represent CG without contour features. The analysis of [kʷ] is shown in (17).

(17) Representing of [kʷ] without contour features

	[kʷ]
Body-[stop]	+
Lips-[round]	+
Glottis-[stiff]	+

In the representation, [+stop] only applies to the articulator Body for [k] and not to Lip of [ʷ]. The gesture [+round] of [ʷ] occurs at the same time as [+stop], which means that the lips are rounded during the closure of [k], which is phonetically true. Finally, there is a single value of [+stiff] (voiceless), which means that [ʷ] is voiceless as well.

Next we consider the representations of more CG units, [kʲ], [tʲ], [pʲ], [pʷ], and [lʷ], shown in (18)-(22).

(18) Feature representation of [kʲ]
Body-[+stop, –fricative]
Tip-[–stop, –fricative]
Glottis-[+stiff]

(19) Feature representation of [tʲ]
Body-[–stop, –fricative]
Tip-[+stop, –fricative]
Glottis-[+stiff]

(20) Feature representation of [pʲ]
Lips-[+stop, –fricative]
Tip-[–stop, –fricative]
Body-[–stop, –fricative]
Glottis-[+stiff]

(21) Feature representation of [pʷ]
Lips-[+stop, –fricative, +round]
Glottis-[+stiff]

(22) Feature representation of [lʷ]
Tip-[–stop, –fricative, +narrow]
Lips-[–stop, –fricative, +round]
Glottis-[–stiff]

Following Padgett (1995), I assume that each articulator can have its own manner features. In [kʲ], Body has the manner features for [k], while the palatal [j] is represented by the co-presence of Tip and Body, where Tip has the manner features for a glide. In [t^j], Tip has the manner features for [t], while [j] is again represented by the co-presence of Tip and Body, where Body has the manner features for a glide. In [p^j], Lips has the manner features for [p], while features under Tip and Body represent the glide [j]. In [p^w], Lips has the manner features for [p], as well as the feature [+round] for [w]. Finally, in [l^w], Tip has the features for [l] and Lips has the features for [w].

The analysis shows that we should not think in terms of IPA symbols, because the same IPA symbol often has different features. For example, while [j] always involves the co-presence of Tip and Body, only one needs to be [–stop, –fricative]: in [k^j] only Tip is [–stop, –fricative], in [t^j] only Body is [–stop, –fricative], and in [p^j] both Tip and Body are [–stop, –fricative]. Similarly, in English word-initial position, [j] is (at least partially) voiceless in [p^j], [t^j], and [k^j] but voiced in [b^j], [d^j], and [g^j] (Ladefoged and Johnson 2011: 73).

7.2.3 CL (consonant–liquid) units

A CL unit is made of a consonant and a liquid, such as [br] and [bl]. A CL unit seems to contain sequential feature values. In traditional features, for example, [bl] seems to call for the representation in (23).

(23) Literal representing of [bl] in traditional features

	[b	l]	
[stop]	+	–	← contour feature
[lateral]	–	+	← contour feature

However, it is possible to represent some CL units without contour features. The analysis of [bl] is shown in (24), where [+narrow] represents a "lateral" sound (Chapter 6).

(24) Feature representation of [bl] without contour features
Lips-[+stop, –fricative]
Tip-[+narrow]
Glottis-[–stiff]

Because there is no contour feature, [bl] is a possible single sound. Similarly, it can be shown that [pr], [pl], [fr], [fl], [kr], [kl], etc., where [r] and [l] are voiceless, are also possible single sounds.

Next, we consider [dr] in English. It may appear that their features are incompatible. For example, [d] is [+stop] and [r] is [–stop], both under Tip. In addition, [d] is not retroflex while [r] is. But consider the feature structure in (25), which does not contain contour features. Since the English [r] is [+round] in the syllable onset, it is thus specified as well.

(25) Feature representation of [dr] (as [ɖʐʷ])
Lips-[+round]
Tip-[+stop, +fricative, –back]
Glottis-[–stiff]

Since there is no contour feature in the representation, the structure is a well-formed single sound. The question is whether the structure is a proper representation of [dr] in English. According to the features, (25) is a rounded retroflex affricate [ɖʐʷ]. It is clear that [dr] is rounded. In addition, it is reasonable to say that [d] is retroflex in [dr], owing to assimilation with [r]. But is [dr] an affricate? The answer seems to be yes, as many linguists have suggested (Jones 1962: 165–6; Abercrombie 1967: 148; Gimson 1970: 171; Lawrence 2000: 86).

Next we consider [dl], which raises the question of whether "stop" in [d] and "lateral" in [l] are compatible. If "stop" is defined as complete closure and "lateral" as leaving the sides of the tongue open, as traditionally defined, then the two feature values are incompatible. On the other hand, if "lateral" is defined as [+narrow] (narrowing the tongue blade; Ladefoged 1980 and Browman and Goldstein 1989), then [+stop] and [+narrow] are compatible. This approach is similar to our analysis of affricates above. The representation of [dl] is shown in (26).

(26) Feature representation of [dl]
Tip-[+stop, +narrow]
Glottis-[–stiff]

In this analysis, the stop is made with a narrowed tongue blade, whose sides are easier to open (than those in a stop or fricative), and so the side opening occurs slightly ahead of the release of the tongue tip. The representation of [tɬ] (sometimes simply written as [tl]) would be similar, except the feature for Glottis would be [+stiff], the gesture for voiceless.

We have seen that some CL units are possible single sounds. However, other CL units have incompatible features. For example, while [fr] is a possible single sound, [sr] and [θr] are not. The analysis of [sr] is shown in (27), where [r] is unspecified for [voice].

(27) Feature representation of [s], [r], and [sr]

[s]	[r]	[sr]
Tip-[–stop, +fricative]	Tip-[–stop, –fricative]	Tip-[–stop, +fricative, –fricative]
Glottis-[+stiff]		Glottis-[+stiff]

In [sr], [s] is Tip-[+fricative] but [r] is Tip-[–fricative], resulting in a contour feature Tip-[+fricative, –fricative]. Similarly, in [θr], [θ] is Tip-[+fricative] but [r] is

Tip-[−fricative], resulting in a contour feature Tip-[+fricative, −fricative]. In addition, [θ] is Tip-[+front] (interdental) but [r] is Tip-[−front] (retroflex), resulting in a contour feature Tip-[+front, −front].

We conclude that many CL units, such as [pl pr pw dr dl], can be represented as single sounds but others, such as [sr θr], cannot and must be represented as two-sound units.

7.3 Complex sounds whose features are incompatible

I discuss two cases of incompatible complex sounds: contour tones, and pre- and post-nasalized stops. I show that both involve contour features and must be represented as two-sound clusters.

7.3.1 Contour tones

A contour tone is a rise or a fall (and sometimes both, as in rise–fall or fall–rise). Contour tones present a problem for feature theory if they occur on a short vowel. For example, if we represent a rise as [−H, +H], where [H] is the tone feature "high," then it is a contour feature, and according to the No Contour Principle it cannot occur on a short vowel. This is shown in (28) and (29).

(28) A contour tone on a short vowel

```
   [a]
   /\
[−H][+H]  ←  contour feature
```

(29) A contour tone on a long vowel (same as two short vowels)

```
 [a     a]
  |     |
[−H]  [+H]  ←  no contour feature
```

Two kinds of contour tone have been proposed, "contour tone units" and "contour tone clusters" (Yip 1989). Unit contour tones and cluster contour tones are phonetically identical. Their difference is determined solely by their behavior. For example, a cluster rise can be created when [−H] is added to the left of [+H], but a unit rise cannot. Similarly, a cluster rise can split into [−H] and [+H] over two syllables, but a unit rise cannot.

Two common arguments for contour tone units have been proposed (Yip 1989; Barrie 2007). First, they do not split into level tones, as just mentioned. Second, they can spread like a unit. For example, if a unit rise is spread over three syllables, we expect to get a sequence of rise–rise–rise, and such a case is reported in Danyang

Chinese. However, both arguments have been questioned (Duanmu 1994). In what follows, therefore, I exclude contour tone units and consider contour tone clusters only.

Many cases of contour tone clusters have been reported. While not all of them can be reviewed here, I shall discuss some well-known cases. In particular, I focus on whether a contour tone can fall on a short syllable. The case in Asian languages is clear, where contour tones fall on long syllables only. In addition, complex contour tones are found on pre-pause syllables only, where such syllables are extra-long (Woo 1969; Bao 1990a; Lin and Yan 1988). In African languages, vowels with contour tones are often long or lengthened as well. This is the case, for example, in Efik (Ward 1933) and Mende (Aginsky 1935; Crosby 1944; Ward 1944). It is worth noting that vowel lengthening is easy to overlook because it is not always transcribed. For illustration, let us consider Mende.

Aginsky (1935), whose description is noted for its "relative exactness" (Spears 1967: 231), consistently transcribes Mende vowels as long when they bear contour tones (what she calls tone "glides"). Similarly, Ward (1944: 3) observes that "vowels on a falling or a rising tone sound long to English ears," and Crosby (1944) transcribes vowels as long when they carry contour tones. Thus, it seems clear that contour tones fall on long vowels. However, in the phonemic transcription of Spears (1967), length is not indicated for vowels that carry contour tones. This is understandable, because phonemic transcription ignores predictable variation, such as vowel lengthening under a contour tone. Similarly, Innes (1969) does not transcribe vowel length either. However, the lack of vowel length transcription is often—erroneously—taken to be the lack of vowel length itself. For example, Leben (1973: 50) adopts the transcription of Spears (1967) and claims that contour tones in Mende fall on short vowels. Leben's work in turn is cited by Goldsmith (1976), van der Hulst and Smith (1982), and many textbooks (Kenstowicz and Kisseberth 1979; Halle and Clements 1983; Katamba 1989; Kenstowicz 1994), all of which claim that contour tones fall on short vowels in Mende.

In summary, despite frequent claims to the contrary, there is no compelling evidence that contour tones can fall on phonetically short vowels. The lack of short contour tones has led to the proposal of the No Contour Principle (Duanmu 1994). The lack of short contour tones is also consistent with the perception study of Greenberg and Zee (1979), who have found that when a short vowel carries a rising F0 contour (e.g. a 50Hz rise over a vowel of 90ms), listeners cannot perceive a contour tone.

7.3.2 *Pre- and post-nasalized consonants*

Pre-nasalized consonants, such as [mb] in the Mende word [mba] 'rice,' present a problem if they are single sounds, because they require a contour feature [+nasal,

–nasal] (Anderson 1976). The same is true for post-nasalized consonants, or what are called "nasally released" consonants in UPSID, such as [bm]. Moreover, contrastive pairs such as [mb]-[mp] and [nd]-[nt] have been reported in Konyagi (Santos 1977) and Sama (Verheijen 1986). Such cases involve not just a contour feature for [nasal] but a contour feature for [voice] as well.

Anderson (1976) argues that pre- and post-nasalized consonants must be recognized as single sounds. In addition, they require contour features. However, Anderson is aware that the use of contour features will lead to over-prediction of possible sounds, most of which are not found. To avoid this, Anderson proposes that contour features be limited to two cases: contour tones and contour [nasal]. Anderson's proposal has two problems. First, there is no explanation why contour features are limited to tones and [nasal]. Second, as discussed above, it is not obvious that contour tone features are needed, because short contour tones have not been clearly established. This would leave contour [nasal] as the only contour feature.

Herbert (1975) distinguishes three cases of pre-nasalized stops, two of which are easy to account for. In the first, a language lacks a four-way contrast among voiceless stops, voiced stops, nasals, and pre-nasalized stops, such as [p b m mb]. If there is only a three-way contrast, such as [p mb m], then [mb] could simply be (a fully voiced) [b], because there is no contrast between [mb] and [b].

The second case is called the "shielding effect," where a short nasal or oral stop appears between an oral sound and a nasal sound, owing to a temporal misalignment between the velum gesture and the oral gesture. For example, in the English word *warmth*, [mθ] is often pronounced as [m^{p}θ], where a short [p] shields the nasal [m] from the oral [θ]. Similarly, in Kaingang (Wiesemann 1972), [ãb] becomes [ã^{m}b], where a short [m] shields a nasal [ã] from [b], and [ma] becomes [m^{b}a], where a short [b] shields [m] from an oral [a]. Under the shielding effect, the occurrence of pre- and post-nasalized consonants is predictable from the nasality of neighboring sounds. In addition, there is no contrast between a pre- or post-nasalized consonant and a regular one. For example, there is no contrast in Kaingang between [ãb] and [ã^{m}b], or between [ma] and [m^{b}a]. Therefore, we do not need to create [^{m}b] or [m^{b}] as new sounds.

Beddor and Onsuwan (2003) describe the phonetic properties of stops in Ikalanga, which is reported to have a four-way contrast among [p b m mb]. When [b m mb] occur between vowels, VbV, VmV, and VmbV are realized as VbV, VmṼ, and Vm^{b}V respectively, where [Ṽ] is a heavily nasalized vowel and [^{b}] is a very short [b] (V is slightly nasalized before [m] and [mb], too, as expected). Beddor and Onsuwan suggest that the cue for [mb] is the lack of nasalization in the following vowel, whereas the cue for [m] is the presence of heavy nasalization on the following vowel. Given such facts, a shielding analysis is also available: Ikalanga has both oral and nasal vowels, where [^{b}] results from the shielding effect between a nasal and an oral vowel. Therefore, there is no need to postulate [mb] as a new phoneme.

In the third case, a pre-nasalized (or post-nasalized) stop contrasts with a single stop or nasal, such as [mb] vs. [p b m]. In this case, we must consider whether the pre-nasalized stop is a single sound or a cluster of two sounds. The argument here often depends on syllable structure. Let us consider the case of Luganda, which is fairly representative.

Consonant distributions in Luganda are shown in (30), where NC is a pre-nasalized consonant and CC is a geminate consonant (Ashton et al. 1954; Tucker 1962; Herbert 1975).

(30) Consonant distributions in Luganda (glides are omitted)

Word-initial:	C or NC
Word-medial:	C, NC, or CC
Word-final:	None

Since no word in Luganda ends in a vowel, two analyses have been proposed. In the first (Ashton et al. 1954), all syllables are open (ending in a vowel); in this analysis, NC is a single consonant, otherwise it would be the only type of consonant cluster. In the second analysis (Tucker 1962; Herbert 1975), a syllable can be CV, CVC, or N, where CVC cannot occur in word-final position. A medial NC is a cluster that is split between two syllables, so is CC. An initial NC is also a cluster, where N is syllabic and forms a syllable by itself. There is some evidence for the second analysis. For example, in Italian, all words end in a vowel, and CVC syllables are found in non-final positions only. Similarly, in Japanese, medial clusters are limited to NC and geminates.

The example shows that, despite reported cases of pre-nasalized (and post-nasalized) consonants, conclusive evidence for them to be single sounds is rare, whereas a cluster analysis is often available. Therefore, there is no compelling reason to consider pre-nasalized and post-nasalized consonants to be possible single sounds, nor is there any need to consider post-nasalized consonants to be "obstruent nasals" (Durvasula 2009). For example, in the cluster analysis, [mb] no longer presents a problem for feature theory, because it is made of two regular sounds, [m] and [b]. Similarly, [mp] and [nt] present no problem either, since [m], [p], [n], and [t] are all regular sounds.

7.4 Non-pulmonic sounds

Non-pulmonic sounds include clicks, ejectives, and implosives. The production of a non-pulmonic sound involves a sequence of steps. Therefore, the feature representation of a non-pulmonic sound has been problematic, since it seems to require contour features. I shall show that, in some contexts, some steps can be realized in the preceding or following sound. Therefore, it is sometimes possible to represent a non-pulmonic sound as a single sound, such as when it occurs between vowels. It is possible, too, to represent a non-pulmonic sound as a two-sound unit. I start with one-sound representations of clicks, ejectives, and implosives, followed by two-sound representations of them.

7.4.1 Clicks

Although clicks are not used as phonemes in many languages, they are easy to make for many people. For example, most English speakers can make the dental click [ǀ], as in *tsk-tsk*.

According to Ladefoged and Johnson (2011: 144), the production of a click involves four sequential steps. Those for [ǀ] are shown in (31) and (32).

(31) Sequential steps in producing the click [ǀ] (Ladefoged and Johnson 2011: 144)
Step 1: Closure of the tongue tip and the tongue body.
Step 2: While maintaining the closures, lower the tongue body.
Step 3: Release the tongue tip.
Step 4: Release the tongue body.

The feature interpretation of the four steps is shown in (32), where I interpret the lowering of the tongue body as the result of retracting the tongue root.

(32) Feature representation of the four steps in producing [ǀ]

	1	2	3	4	
Tip-[stop]	+	+	–	–	← contour feature
Body-[stop]	+	+	+	–	← contour feature
Root-[advanced]	+	–			← contour feature

In the feature representation, there are three contour features, which present a problem for feature theory. To avoid the problem, some special feature terms have been proposed, such as [suction] (Chomsky and Halle 1968; Halle 2003), [click] (Ladefoged 2007), and [lingual] (Miller 2011). However, such terms are contour features in disguise.

I agree with Ladefoged and Johnson (2011) that the production of a click involves multiple steps. But unlike Ladefoged and Johnson (2011), who assume four steps, I propose that three are sufficient. This is shown in (33).

(33) A three-step feature representation of the production of [ǀ]

	1	2	3
Tip-[stop]		+	–
Body-[stop]		+	(–)
Root-[advanced]	+	–	

In the three-step representation, the closure of Tip and Body need not occur before the retraction of Root but can occur in the same step as the latter. In addition, Body-[–stop] (velar release) is not required but optional. This agrees with the fact that in producing a sequence of clicks, such as [|]-[|]-[|] . . . , there is no velar release between clicks. Moreover, when Tip and Body are both required to release, such as when a vowel follows the click, the releases can occur in the same step. Because of the low oral pressure created by step 2, air will rush into the mouth, creating the click noise, even if both Tip and Body release in step 3.

Now the three steps in producing a click need not all be realized in the click itself. Instead, some of the gestures can be realized in the preceding or following sound. Consider the representation of [a|a], shown in (34), where some features for [a] are omitted.

(34) Feature representation of [a|a]

	[a	|	a]
Tip-[stop]	–	+	–
Body-[stop]	–	+	–
Root-[advanced]	+	–	

In this representation, Root-[+advanced] is realized in the preceding [a] and the release of Tip and Body is realized in the following [a]. The click itself only contains one step, which includes the closure of Tip and Body and the retraction of Root, and no contour feature is involved.

Next we consider the analysis of [. . . | | | . . .], a sequence of [|] with no velar release. The representation is shown in (35).

(35) Feature representation of [. . . | | | . . .]

	[. . .	|	k˺	|	k˺	|	. . .]
Tip-[stop]		+	–	+	–	+	–
Body-[stop]		+	+	+	+	+	
Root-[advanced]	+	–	+	–	+	–	

Each [|] is preceded by Root-[+advanced] and followed by Tip-[-stop], in agreement with the three-step analysis. However, the click itself only contains one step. The step in between clicks has an unreleased velar closure, which is equivalent to [k˺] (an unreleased [k]), which serves the dual purpose of realizing the third step of the preceding click and the first step of the following click. In other words, [. . . | | | . . .] is equivalent to [. . . | k˺ | k˺ | . . .].

Clicks can also involve additional features, such as [!n] (a nasalized [!]) and [!g] (a voiced [!]) in !Xóõ. Since a regular click only involves the oral cavity, while the glottis and the velum are free to add additional gestures, such variations of clicks pose no problem for feature theory. Therefore, they are not discussed here.

7.4.2 Ejectives

Ladefoged and Johnson (2011: 137) consider there to be four sequential steps in the production of ejectives. The analysis of [k'] is shown in (36) and its feature representation is shown in (37).

(36) Steps in producing the ejective [k'] (Ladefoged and Johnson 2011: 137)
Step 1: Closure of the tongue body and the glottis.
Step 2: While maintaining the closures, raise the larynx.
Step 3: Release the tongue closure.
Step 4: Release glottal closure.

(37) Feature representation of the four steps in producing [k']

	1	2	3	4	
Body-[stop]	+	+	–	–	← contour feature
Larynx-[raised]	–	+	+	+	← contour feature
Glottis-[stop]	+	+	+	–	← contour feature

The four-step representation again involves contour features. To avoid the problem, a special feature value was proposed, such as [+ejection] (Chomsky and Halle 1968: 323). A similar proposal was offered by Ladefoged (2007: 162). However, such proposals in effect assume contour features.

I agree with Ladefoged and Johnson (2011) that a complete ejective event involves a sequence of steps, and that the build-up of oral pressure is provided by raising the larynx. However, instead of four steps, I propose that three steps are sufficient. The analysis of [k'] is shown in (38).

(38) A three-step representation of the production of [k']

	1	2	3
Body-[stop]		+	–
Larynx-[raised]	–	+	
Glottis-[stop]		+	(–)

First, the closure of Body and Glottis need not precede the raising of Larynx but can occur in the same step as the latter. Second, the opening of Glottis is not required but optional. Third, when Body release and Glottis release are both required (such as when a vowel follows), they can occur in the same step. The simultaneous release of Body and Glottis will still cause air to flow outward, if the next sound is pulmonic, in which the lung pressure is higher than the outside air pressure.

The three steps need not be all realized in the ejective itself. Instead, the first step can be realized in the preceding sound and the third in the following sound. This can be seen in the analysis of [ak'a], shown in (39), where some features for [a] are omitted.

(39) Analysis of [ak'a]

	[a	k'	a]
Body-[stop]	–	+	–
Larynx-[raised]	–	+	
Glottis-[stop]	–	+	–

In the analysis, Larynx-[–raised] is realized in the preceding [a], the release of Body and Glottis is realized in the following [a], and [k'] itself contains the second step, with no contour feature involved.

The idea of not including the release of a stop from feature representation is not new. For example, when [p] is pronounced alone, it seems to involve two steps, a closure and a release, but in [pa], the release phase is provided by [a]. Therefore, we do not consider the pronunciation of [p] to involve two separate steps.

In summary, ejectives can be represented as single sounds, provided they occur with other sounds that provide some relevant features to complete the full ejective event.

7.4.3 Implosives

According to Ladefoged and Johnson (2011: 141), the production of implosives involves three sequential steps. The analysis of [ɓ] is shown in (40) and (41).

(40) Steps in producing the implosive [ɓ] (Ladefoged and Johnson 2011: 141)
Step 1: Closure of the lips (Labial).
Step 2: While maintaining the closure, lower the larynx.
Step 3: Release the labial closure.

(41) Feature representation of [ɓ]

	1	2	3	
Lips-[stop]	+	+	–	← contour feature
Larynx-[raised]	+	–		← contour feature

Like Ladefoged and Johnson (2011), I assume that a complete implosive event for [ɓ] involves three steps, although Lip closure need not start in step 1 but can start in step 2. Now if all the three steps are realized in an implosive, there will be contour features. To avoid the problem, special feature terms have been proposed, such as [suction] (Chomsky and Halle 1968) and [implosive] (Ladefoged 2007: 162). However, such terms are contour features in disguise.

If instead some features are realized on the surrounding sounds, then the implosive itself may involve just one step, without contour features. Specifically, Larynx raising can be realized in the preceding sound and Lip release can be realized in the following sound. Therefore, when [ɓ] occurs between vowels, only one step is realized in [ɓ] itself. The analysis of [aɓa] is shown in (42), where some features of [a] are omitted.

(42) Feature representation of [aɓa]

	[a	ɓ	a]
Lips-[stop]	–	+	–
Larynx-[raised]	+	–	

Implosives are often voiced. In some languages both voiced and voiceless implosives are reported, such as [ɓ ɓ̥] in Ngiti (Kutsch Lojenga 1994), where "voiceless" implosives probably involve either a creaky voice or "a full glottal closure" (Ladefoged and Maddieson 1996: 87). In the present analysis, the basic components of an implosive are an oral closure and the lowering of the larynx, while the glottis is free to assume a gesture for regular voice, creaky voice, or glottal stop. Therefore, both kinds of implosives are in principle possible, too.

7.4.4 Two-sound analysis

In the preceding discussion, I have shown that it is possible to represent clicks, ejectives, and implosives as single sounds. It is also possible to represent them as two-sound units, which are closer to traditional phonetic descriptions. Examples of single-sound and two-sound representations of clicks, ejectives, and implosives are shown in (43)–(45) respectively, where voicing features are omitted.

(43) Feature representation of [aǀa]

	[ǀ] as one sound			[ǀ] as two sounds			
	[a	ǀ	a]	[a	ǀ	ǀ	a]
Tip-[stop]	–	+	–	–	+	+	–
Body-[stop]	–	+	–	–	+	+	–
Root-[advanced]	+	–			+	–	

(44) Feature representation of [ak'a]

	[k'] as one sound			[k'] as two sounds			
	[a	k'	a]	[a	k'	k'	a]
Body-[stop]	–	+	–	–	+	+	–
Larynx-[raised]	–	+			–	+	
Glottis-[stop]	–	+	–	–	+	+	–

(45) Feature representation of [aɓa]

	[ɓ] as one sound			[ɓ] as two sounds			
	[a	ɓ	a]	[a	ɓ	ɓ	a]
Lips-[stop]	–	+	–	–	+	+	–
Larynx-[raised]	+	–			+	–	

The single-sound representation and the two-sound representation make different predictions both phonetically and phonologically. Phonetically, the single-sound representation predicts that a non-pulmonic sound has a similar duration to a pulmonic one. There is partial evidence for this. For example, it has been reported that implosives have similar (or slightly shorter) durations compared to regular voiced stops (Nihalani 1986), although durational data that compare regular stops and other non-pulmonic sounds are hard to find.

Phonologically, both the single-sound and the two-sound representations yield good syllable structures. Specifically, an intervocalic single-sound representation can be syllabified as V.CV or VC.V, both yielding simple syllables without consonant clusters. An intervocalic two-sound representation can be syllabified as VC.CV, which yields simple syllables too without consonant clusters. However, the two-sound representation predicts that non-pulmonic sounds are highly restricted otherwise. For example, in intervocalic position, a non-pulmonic sound ought not to occur with another stop, otherwise we would get VC.<C>.CV, where an unsyllabified <C> will be left in the middle, assuming that neither the onset nor the coda can contain two stops. Since there is a lack of relevant phonological data, I leave it open as to whether the single-sound or the two-sound representation is more appropriate.

7.5 Summary

I have discussed a set of cases, loosely referred to as "complex sounds," that have posed problems for feature theory. I have shown that the No Contour Principle offers a precise definition of feature compatibility and of what is or is not a possible complex sound. A possible complex sound is one that can be represented as a single sound, such as homorganic affricates [ts] and [kx], non-homorganic affricates [ps] and [ks], CG (consonant–glide) units [kw] and [kj], and CL (consonant–liquid) units [pl] and [pr]. In addition, non-pulmonic sounds (clicks, ejectives, and implosives), which are either left out of a feature theory (e.g. Browman and Goldstein 1989) or represented with special features, such as [suction], [click], [ejection], and [implosive] (Chomsky and Halle 1968; Halle 1995; Ladefoged 2007), can also be represented as single sounds when they occur between vowels, and no special feature is needed.

I have also shown that some complex sounds, in particular contour tones, prenasalized stops, and post-nasalized stops, cannot be represented as single sounds. Instead, they should be represented as two-sound clusters, and evidence is offered that they are.

The analysis of complex sounds offers further support for the feature system proposed in Chapter 6, which is strikingly simple, yet powerful enough to generate a sufficient number of well-formed structures, while restrictive enough to minimize over-prediction.

8

Concluding remarks

The goal of this study is to determine a feature system that is minimally sufficient to distinguish all consonants and vowels in the world's languages, based on close examinations of two databases of transcribed sound inventories, UPSID (451 inventories) and P-base (628 inventories). No prior theoretical assumption was made as to what properties the feature system should have, such as whether features should be innate or binary.

The same goal has been explored in many previous feature theories, such as Trubetzkoy (1939), Jakobson et al. 1952, Chomsky and Halle (1968), Ladefoged (1971; 2007), Clements (1985), Sagey (1986), Clements and Hume (1995), and Halle (1995; 2003). Some of them have also used data from many languages. For example, Trubetzkoy (1939) cited some 200 languages, Jakobson et al. (1952) cited nearly seventy languages, and Maddieson (1984) used a database of 318 languages. So what is new in the present study?

The most important contribution of the present study, in my view, is a solution to a methodological problem in interpreting data from different languages. Specifically, let X be a sound from one language and Y a sound from another language. How do we decide whether X and Y should be treated as the same sound or different sounds? It is well known that, when X and Y are transcribed with the same phonetic symbol, there is no guarantee that they are phonetically the same. Similarly, when X and Y are transcribed with different phonetic symbols, there is no guarantee that they must be different sounds. Moreover, even if X and Y are somewhat different phonetically, there is no guarantee that they cannot be treated as the same sound in a language. Without a solution to the methodological problem, little use can be made of databases like UPSID and P-base. Indeed, some linguists believe that cross-language comparison of speech sounds is impossible, in that no sound in one language can meaningfully be identified with one in another (e.g. Boas 1896; Joos 1957; Ladefoged 1972, 1992; Disner 1983; Port and Leary 2005).

A similar problem exists when we examine just one language, but it does have a solution. Given two sounds A and B in language L, how do we know whether they are the same or different to native speakers of L? The standard solution relies on the notion of contrast, namely, whether A and B can distinguish words. This is shown

A Theory of Phonological Features. First edition. San Duanmu.

in (1), where a phoneme is an abstract category of sameness and an allophone is a concrete realization of a phoneme.

(1) Using contrast to determine phonemes and allophones:
Let A and B (which are not phonetically identical) be two sounds in a language L. If A and B can contrast in L, A and B belong to different phonemes in L. If A and B never contrast in L and if A and B are phonetically similar, A and B are allophones of the same phoneme in L.

For example, [l] and [r] belong to different phonemes in English, because they can distinguish word pairs such *rice–lice*. On the other hand, [l] and [r] are allophones of the same phoneme in Japanese, in which they do not distinguish words.

Now if we extend the notion of contrast from the analysis of one language to the analysis of all languages, a solution to the problem of cross-language comparisons becomes available. Specifically, I have proposed the Principle of Contrast in Chapter 2, repeated here in (2).

(2) Principle of Contrast:
   a. If two sounds A and B can contrast in any language, they must be distinguished by at least one feature.
   b. If two sounds A and B never contrast in any language, they need not be distinguished by a feature.

The proposal ought not to be controversial. Let us consider three cases. First, A and B contrast in every language; in this case, A and B must be distinguished, both in the traditional analysis and in the present one. Second, A and B contrast in some languages and occur as allophones in others; in this case, A and B are still different sounds, both in the traditional analysis and in the present one, in the sense that A and B are represented by different phonetic symbols or different features. In the third case, A and B are different in some way but not identical and never contrast in any languages. In the traditional analysis, A and B are sometimes distinguished as allophones, although not always. In the present analysis, A and B need not be distinguished.

While the Principle of Contrast may seem straightforward, its implementation can be tedious. Let us consider an example. In P-base, we find [i] in English and [i̞] (a "lowered" [i]) in Haitian Creole, but neither language has both. Should [i] and [i̞] be treated as the same sound or as different sounds? To find the answer, we must search through all sound inventories to find out whether [i] and [i̞] ever contrast in any language. We find Tangale, whose vowel inventory is [i i̞ u u̞ e e̞ o o̞ a], where [i] and [i̞] contrast. Does this justify recognizing [i̞] as a new sound (beside [i])? The answer is no, because there is a more common sound [ɪ], which is also lower than [i] and which is not in the vowel inventory of Tangale. Therefore, we have to ask whether [i̞] ever contrasts with [ɪ] in any language. This calls for another search, which yields no hit. Therefore, following the Principle of Contrast, we conclude that [i̞] and [ɪ] can be treated as the same sound and there is no need to recognize [i̞] as a separate sound.

The discussion above has only dealt with one pair of transcriptions. With some 1,000 distinct transcriptions in P-base, it would take a lot of work to check every pair of possible contrasts. Therefore, this study took a different approach. Rather than checking transcribed sounds pair by pair, we check the number of phonetic dimension (or features) we need and the maximal number of contrasts in each dimension. For example, with regard to vowels, we first gather all vowel transcriptions and examine how many phonetic dimensions are involved, such as the backness of the tongue, the height of the tongue, lip rounding, glottalization, nasalization, and so on. Then we examine each pair of features that seem to overlap and check whether both are needed, such as "pharyngealized" vs. [−ATR], or "glottalized" vs. "creaky." We do so by searching through all inventories for such contrasts, such as [iˤ] vs. [ɪ], or [eʔ] vs. [ḛ]. After excluding non-contrastive phonetic dimensions, we then search for the maximal number of contrasts in each dimension and exclude transcriptions that are not contrastive.

The procedure yields a minimally sufficient feature system that can represent all occurring contrasts. The system is simpler than those in the literature, in that it requires fewer features and fewer contrasts in each feature. For example, in the present system, a binary contrast is sufficient in every feature dimension, including vowel height (Chapter 4). In contrast, even the most parsimonious feature theories in the literature, such as Jakobson et al. (1952), Chomsky and Halle (1968), Kiparsky (1974), and Halle (2003), assume at least three degrees of vowel height.

In Chapter 6, the feature system is given a reinterpretation in which features are gestures of articulators. In addition, we have discussed feature specification. Moreover, given the fact that there are so few contrastive gestures, there is a considerable amount of freedom in the realization of a gesture. Therefore, it is not a surprise, but expected, that a sound with the same feature specification can be realized in somewhat different ways phonetically, such as [i] in German and [i] in Norwegian, as observed by Disner (1983). In addition, since such differences are not contrastive in any language, they need not be represented by features (Chapter 6, section 6.7).

The present study has also offered a simple definition of what a sound is (Chapter 1), repeated in (3), based on time and feature compatibility.

(3) Sounds defined by time
    A sound is a set of compatible feature values in one time unit.

Features are compatible if they do not violate the No Contour Principle (Duanmu 1994), discussed in Chapter 7 and rephrased in (4).

(4) No Contour Principle: The structures below contain a contour feature and are ill-formed (X is a time unit; [+F] and [–F] are opposite values of the same feature)

```
      X          or          X
      |                      |
 Articulator            Articulator
     /  \                   /  \
  [+F] [–F]              [–F] [+F]      ← contour feature
```

The No Contour Principle also offers a more constrained definition of possible and impossible complex sounds, without assuming contour features (Chapter 7). Specifically, possible complex sounds include affricates (both homorganic, such as [ts], and non-homorganic, such as [ps]), consonant–glide units (such as [t^j] and [s^w]), consonant–liquid units (such as [pr] and [pl]), and possibly ejectives, clicks, and implosives. Impossible complex sounds include diphthongs (such as [ai]), prenasalized stops (such as [mb]), and contour tones, each of which can, however, be represented as a two-sound cluster.

It is often assumed that, besides representing contrasts, features should represent sound classes ("natural classes") as well (e.g. Halle 1995; Mielke 2008). In this book, little discussion is given on sound classes, not because I think it is unimportant, but because of two other reasons. First, a contrast-based feature system is an independent topic: all contrasts must be represented, regardless of whether they play a role in sound classes. Second, a systematic study of sound classes has just begun and much research is still needed. For example, there was no comprehensive compilation of sound classes until Mielke (2008). In addition, theoretical questions remain as to which patterns imply a sound class. For example, is the set of sounds that can occur after word-initial [s] in English a sound class? Is the set of English consonants that can occur in word-initial position before [l] a sound class? Such questions are not easy to answer and have rarely been asked in the first place. Nevertheless, I hope we can find out, in the not so distant future, whether features based on contrast and those based on sound classes indeed match each other.

I would like to end this chapter with a discussion of three additional questions: How many distinct sounds are there in the world's languages? How many sounds does a language need? Where do features come from?

8.1 How many distinct sounds are there?

There are three ways of counting sounds, shown in (5).

(5) Three ways of counting sounds
   a. Distinct sounds in phoneme inventories
   b. Distinct occurring sounds
   c. Distinct possible sounds for a feature system

First, we can count distinct sounds in the phoneme inventories of the world's languages, with appropriate adjustments; for example, we should exclude diphthongs (Chapter 1) and non-contrastive transcriptions, such as [i̞], which does not contrast with the more common [ɪ]. If we have large databases of phoneme inventories, such as UPSID and P-base, we can arrive at a fairly accurate count. It is true that UPSID and P-base only contain a portion of the world's languages, but it is likely that most phonemes in undocumented languages are similar to those in the documented ones.

Second, we can count occurring sounds, many of which are not represented by the phonemes of a language. For example, in Standard Chinese the syllable onset has a single position, which can be filled by a consonant C, a glide G, or a consonant plus a glide CG, where CG merge into a single complex sound, such as [kʲ], [kʷ], and [ʂʷ] (Duanmu 2007). Now, members of C are found in the phoneme inventory and members of G can be treated as variants of high vowels (Duanmu 2007). However, the set of CG units, which number twice as many as members of C and G combined, are not represented as phonemes in any study. How then do we count occurring sounds like these? Since there is no database of occurring sounds in the world's languages, it is impossible to count them accurately. However, many occurring sounds of a language are found in the phoneme inventories of other languages, if not yet in its own phoneme inventory. For example, Chinese [kʲ] and [kʷ] are found in the phoneme inventory of Hausa, although Chinese [ʂʷ] (as in [ʂʷa] 'brush') is not found in any phoneme inventory in UPSID or P-base.

Third, we can count possible sounds predicted by a given feature system, considering all possibilities of combination, including all possible complex sounds. A similar point is made by Walker and Pullum (1999), who argue that any phonetically pronounceable sound is a phonologically possible sound. While they did not make it clear, Walker and Pullum (1999) must assume a feature system; otherwise there is no way to count sounds. For example, in the present feature system, changing the backness of the tongue can only create a two-way distinction, regardless of how many ways one can position the tongue on the backness scale.

Let us now consider how many distinct vowels there are. We start with vowel phonemes in UPSID. There are 268 distinct vowel transcriptions. If we exclude diphthongs (which total eighty-eight) and long vowels (which total forty), there are 140 vowels left. Among them, many do not contrast, such as [o+] and [o]. If we focus on basic vowels (those involving tongue gesture and lip rounding only), then there are thirty-eight. However, only sixteen basic vowels have been confirmed (Chapter 4). This means that the 140 reported vowels in UPSID can be reduced by half. In (6) I summarize the result for UPSID. The analysis of P-base is similar and is not repeated.

(6) Counting vowels in UPSID (* indicates approximate number)

Diphthongs and long vowels	128
Unconfirmed others	70*
Confirmed distinct vowels	70*
Total distinct transcriptions	268

Next we consider occurring vowels, using Standard Chinese as an example. Standard Chinese has a diminutive suffix that adds a feature [+retroflex] to the vowel, which creates three retroflex vowels [ɚ a˞ u˞] (Duanmu 2007). Since such vowels occur in restricted environments, they are not usually included in the phonemic inventory. Now [ɚ] is found in some phoneme inventories, such as Gelao and Naxi, but [a˞ u˞] are not found in any other language in UPSID or P-base. Therefore, some occurring vowels will be missed if we only count transcriptions in databases like UPSID and P-base.

Finally, we consider possible vowels predicted by feature theory. According to the present study, there are only sixteen basic vowels, defined by four binary features [back, round, high, ATR] (Chapter 4). To the basic set we can add three additional features "nasal," "breathy," and "creaky," creating five additional sets: adding "nasal" only, adding "breathy" only, adding "creaky" only, adding "nasal" and "creaky", and adding "nasal" and "breathy" (assuming that "creaky" and "breathy" cannot combine). This yields a total of $16 \times 6 = 96$ distinct vowels, summarized in (7).

(7) Possible vowels predicted by the present feature system (Chapter 4), by adding "nasal," "breathy," and "creaky" to basic vowels.

Basic vowels	16
Nasalized, breathy, or creaky	48 (=16 × 3)
Nasalized and creaky	16
Nasalized and breathy	16
Total distinct vowels	96

As can be seen, the number of vowels predicted from feature combinations is larger than what is obtained from counting vowel phonemes (i.e. 96 vs. 70). In fact, the total number of possible vowels predicted by feature theory is much larger still. For example, the feature [retroflex], a gesture of the tongue tip, can be added to vowels too, at least to all back vowels. This would add at least 48 more vowels, just three of which, transcribed as [ɨ˞ ɚ a˞], are found in UPSID and P-base.

The counting of consonants is similar. UPSID lists over 600 distinct consonant transcriptions and P-base lists over 900. Some of them can be excluded, such as long consonants and syllabic consonants, which can be represented by the positions they occupy in a syllable. Some are likely to be two-sound units, such as [nd] in Ganda (Herbert 1975) and [ʘx] in !Xóõ (Traill 1985). The confirmed list of consonants will still run to several hundred, though. For illustration, let us consider affricates. If we consider three places of articulation, there are nine possible affricates, both homorganic and non-homorganic, shown in (8).

(8) Predicted and actual affricates in three places of articulation (* indicates not found)

	Labial	Coronal	Dorsal
Labial	[pɸ]	[ps]	[px]*
Coronal	[tɸ]	[ts]	[tx]
Dorsal	[kɸ]*/[kf]	[ks]*	[kx]

Of the nine possible combinations, seven are found in UPSID or P-base, and only [px] and [ks] are not. This appears to show a reasonably close match between sounds predicted by features and sounds found in phoneme inventories. However, let us add another feature to the list, which is [+round]. The result is shown in (9).

(9) Rounded affricates in three places of articulation (* indicates not found)

	Labial	Coronal	Dorsal
Labial	[pɸ^{w}]*	[ps^{w}]*	[px^{w}]*
Coronal	[tɸ^{w}]*	[ts^{w}]*	[tx^{w}]*
Dorsal	[kɸ^{w}]*/[kf^{w}]*	[ks^{w}]*	[kx^{w}]*

This time, none of the predicted affricates is found in UPSID or P-base, even though there is no obvious problem in rounded affricates. For example, [ps], [p^{w}], and [s^{w}] are all occurring sounds, which means that there ought to be no problem of compatibility among [p], [s], and [w], and that [ps^{w}] ought to be well-formed as far as feature structure is concerned.

Finally, let us consider rounded fricatives in P-base, shown in (10). Once again, for most of the items, there is no problem in feature compatibility, yet many predicted sounds are not found.

(10) Predicted and occurring rounded fricatives in P-base (* indicates not found)

[ɸ^{w}]*	[f^{w}]	[θ^{w}]*	[s^{w}]	[ʂ^{w}]*	[ʃ^{w}]	[ɕ^{w}]*	[ç^{w}]	[x^{w}]	[χ^{w}]	[ħ^{w}]	[h^{w}]
[β^{w}]*	[v^{w}]	[ð^{w}]*	[z^{w}]	[ʐ^{w}]*	[ʒ^{w}]	[ʑ^{w}]*	[ʝ^{w}]*	[ɣ^{w}]	[ʁ^{w}]	[ʕ^{w}]	[ɦ^{w}]

What we see in consonants is similar to what we see in vowels. In both cases, feature theory predicts far more possible sounds than are found in phoneme inventories, at least for non-basic consonants and vowels. This raises a question: Why are so many possible sounds not used?

8.2 How many sounds does a language need?

We have seen that, even though the present feature system is simpler than previous ones, it still predicts many more sounds than are found in the phoneme inventories of the world's languages. This discrepancy calls for an explanation.

One might suggest that there are several times more undocumented languages than are included in UPSID and P-base, and many of the missing sounds may be found in them. This scenario is unlikely, though, given the rarity of "surprise" languages and given how large the number of missing sounds is.

It would be useful, instead, to ask how many sounds a language needs. If a language does not need many sounds, then it would be understandable why complex sounds are rarely chosen. To begin, let us assume that we want to distinguish all morphemes in a language (where morphemes are meaningful pieces that build words). Next, let us ask how many morphemes a language has. According to my calculation, the English lexicon CELEX (Baayen et al. 1995) has about 10,000 morphemes, if we exclude "zero derivations." For example, *study* (verb) counts as a morpheme but *study* (noun) does not, because the latter is thought to have a zero suffix. This way of counting is reasonable, because words of different parts of speech (e.g. verb vs. noun) are unlikely to cause ambiguity in context. Now by the same method of counting, Chinese also has about 10,000 morphemes (Huang and Duanmu 2013), if we exclude word pairs that only differ in part of speech; for example, we count *xuexi* 'study' (verb) and *xuexi* 'study' (noun) just once, not twice. Given the fact that English and Chinese are two of the largest languages in the world, it is unlikely that any other language has more than 10,000 morphemes.

Next we consider three notions: morpheme form (the consonants and vowels it contains), form count (how many distinct forms there are), and homophone rate. If we use the average sizes of consonant and vowel inventories in P-base, the result is shown in (11).

(11) Morpheme form, form count, and homophone rate, assuming an average phoneme inventory size of eight vowels and twenty-two consonants

Morpheme count	*Morpheme form*	*Form count*	*Homophone rate*
10,000	CVC	3,872	10,000/3,872 = 2.6
10,000	CVCV	30,976	10,000/30,976 = 0.3

If the average morpheme has the form CVC, there are 3,872 distinct forms, and the homophone rate is 2.6. If the average morpheme form is CVCV, then there are three times as many distinct morpheme forms as needed.

The example shows that, given a limited number of morphemes to distinguish, a language can easily come up with enough morpheme forms without using many consonants and vowels. Therefore, it is natural that many possible sounds are not used, especially those that are more difficult, involving multiple features or articulators.

8.3 Where do features come from?

Let us close this chapter by considering a question that has generated much interest and debate: Where do features come from? Let us consider three well-known proposals (see Clements and Ridouane 2011 for more views).

Chomsky and Halle (1968: 4) propose that features are "substantive universals," mentally innate for all humans. This proposal intends to explain why a small set of features seem to recur in the world's languages. If every language comes up with its own means to distinguish sounds, we should perhaps expect there to be a lot more variation across languages.

Stevens (1972; 1989) proposes that features are physically universal, but not mentally so. In particular, as an articulator moves along a phonetic dimension (such as the backness of the tongue), the resulting acoustic change is smooth in some regions but "quantal" in others. A quantal change is one where a small movement of the articulator leads to a critical change in acoustic signal. Gestures located on different sides of a quantal change are ideal for phonemic contrast, because the acoustic difference is large and robust. Gestures located on the same side of a quantal change are unsuitable for phonemic contrast, because the acoustic difference is small. Gestures located too close to a quantal change are also unstable for phonemic contrast, because the acoustic result is unstable.

Ladefoged (1972) proposes yet a different view, according to which features are phonetic scales. The set of scales are physiologically determined and available to all languages, but each language is free to choose which scales to make use of, how many contrastive degrees a scale is divided into, and where the dividing lines are. On this view, feature scales are universal, but feature values are not, and nothing is mentally innate (apart from a general reasoning ability).

The present study confirms some predictions of the innateness proposal: (i) the number of features is small, (ii) all features are binary, and (iii) features can be compared across languages. Indeed, the present study yields a feature system that is simpler than the innateness proposal envisioned. Still, the innateness proposal remains a hypothesis, and it is not the only way to explain the simplicity of the feature system. In particular, the presence of features (as phonetic dimensions or scales) could be physiologically determined but not mentally so. In addition, the use of binary feature values can be attributed to two non-mental factors: (i) a language does not need many phonemic contrasts, as discussed in the previous section, and (ii) from a functional point of view, binary contrasts are more robust, whereas multi-level contrasts are too complicated to be adopted by a speech community, especially when simpler solutions are readily available.

The quantal proposal remains a hypothesis, too. In addition, it is unclear how many contrasts it predicts for each feature (phonetic dimension). Moreover, assuming that different articulatory gestures can produce different acoustic effects, there is no need to assume that every feature or gesture must rely on a quantal change.

The proposal of phonetic scales also fails to predict how many degrees of contrast a language can make use of in each feature; in fact, the proposal assumes no theoretical limit in the first place. In addition, the proposal either overlooks or rejects a mapping relation between feature values across languages—a relation the present study has shown to be possible.

It is a remarkable discovery that all sounds in the world's languages can be represented by a small set of features. Indeed, two leading scholars, Ladefoged and Halle (1988: 577), consider it to be "the most fundamental insight gained during the last century." In addition, feature theory has remained "part of the heart of phonology", even though it has occasionally been set aside (Rice 2003: 404). Still, there is much room to improve, and I hope the present study has made new progress.

References

Abercrombie, David (1967). *Elements of General Phonetics*. Chicago: Aldine.

Aginsky, Ethel (1935). *A Grammar of the Mende Language*. Philadelphia: Linguistic Society of America.

Allen, Blake, Douglas Pulleyblank, and Ọ̀ládiípọ̀ Ajíbóyè (2013). Articulatory mapping of Yoruba vowels: an ultrasound study. *Phonology* 30: 183–210.

Allen, E. J. (1976). Dizi. In M. Lionel Bender (ed.), *The Non-Semitic Languages of Ethiopia*, 377–92. East Lansing: African Studies Center, Michigan State University.

Anceaux, Johannes Cornelis (1965). *The Nimboran Language: Phonology and Morphology*. The Hague: Nijhoff.

Anderson, Stephen R. (1974). *The Organization of Phonology*. New York: Academic Press.

Anderson, Stephen R. (1976). Nasal consonants and the internal structure of segments. *Language* 52(2): 326–44.

Ao, Benjamin X. P. (1993). Phonetics and phonology of Nantong Chinese. Doctoral dissertation, Ohio State University, Columbus.

Archangeli, Diana (1988). Aspects of underspecification theory. *Phonology* 5: 183–207.

Archangeli, Diana (2011). Feature specification and underspecification. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds), *The Blackwell Companion to Phonology*, vol. 1, 148–70. Oxford: Wiley-Blackwell.

Aronson, Howard I. (1990). *Georgian: A Reading Grammar*. Columbus, OH: Slavica.

Aronson, Howard I. (1997). Georgian phonology. In Alan S. Kaye (ed.), *Phonologies of Asia and Africa*, vol. 2, 929–39. Winona Lake, Ind.: Eisenbrauns.

Ashton, E. O., E. M. K. Mulira, E. G. M. Ndawula, and A. N. Tucker (1954). *A Luganda Grammar*. London: Longmans.

Baayen, R. Harald, Richard Piepenbrock, and L. Gulikers (1995). *The CELEX Lexical Database: Release 2* (CD-ROM). Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bailey, Charles James Nice (1978). *Gradience in English Syllabification and a Revised Concept of Unmarked Syllabification*. Bloomington: Indiana University Linguistics Club.

Bao, Zhiming (1990a). On the nature of tone. Doctoral dissertation, MIT.

Bao, Zhiming (1990b). Fanqie languages and reduplication. *Linguistic Inquiry* 21(3): 317–50.

Bao, Zhiming (1999). *The Structure of Tone*. New York: Oxford University Press.

Barrie, Michael (2007). Contour tones and contrast in Chinese languages. *Journal of East Asian Linguistics* 16(4): 337–62.

Basbøll, Hans (1981). Remarks on distinctive features and markedness in generative phonology. In Andriana Belletti, Luciana Brandi, and Luigi Tizzi (eds), *Theory of Markedness in Generative Phonology: Proceedings of the 1979 GLOW Conference*, 25–64. Pisa: Scuola normale superiore.

Bearth, Thomas, and Hugo Zemp (1967). The phonology of Dan (Santa). *Journal of African Languages* 6: 9–29.

Beckman, Jill, Michael Jessen, and Catherine Ringen (2009). German fricatives: coda devoicing or positional faithfulness? *Phonology* 26(2): 231–68.

Beckman, Mary, and Janet Pierrehumbert (1986). Intonation structure in Japanese and English. *Phonology Yearbook* 3: 255–309.

Beddor, Patrice Speeter, James D. Harnsberger, and Stephanie Lindemann (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics* 30.4: 591–627.

Beddor, Patrice Speeter, and Chutamanee Onsuwan (2003). Perception of prenasalized stops. Paper presented at the 15th International Congress of Phonetic Sciences (ICPhS 2003), Barcelona.

Benua, Laura (1995). Identity effects in morphological truncation. In Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk (eds), *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, 77–136. Amherst, Mass.: Graduate Linguistic Student Association.

Black, Mr and Mrs K. (1971). *The Moro Language: Grammar and Dictionary*. Khartoum: Sudan Research Unit, University of Khartoum.

Boas, Franz (1896). The limitations of the comparative method of anthropology. *Science* n.s. 4(103), 901–8.

Boersma, Paul (2001). Praat, a system for doing phonetics by computer. *Glot International* 5(9/10): 341–5.

Bornstein, Marc H., and Helen G. Bornstein (1976). The pace of life. *Nature* 259(5544): 557–9.

Borowsky, Toni (1986). Topics in the lexical phonology of English. Doctoral dissertation, University of Massachusetts, Amherst.

Bradshaw, Mary (1999). A crosslinguistic study of consonant–tone interaction. Doctoral dissertation, Ohio State University, Columbus.

Brakel, Arthur (1983). *Phonological Markedness and Distinctive Features*. Bloomington: Indiana University Press.

Browman, Catherine P., and Louis M. Goldstein (1986). Towards an articulatory phonology. *Phonology Yearbook* 3: 219–52.

Browman, Catherine P., and Louis M. Goldstein (1989). Articulatory gestures as phonological units. *Phonology* 6: 201–51.

Calabrese, Andrea (1995). A constraint-based theory of phonological markedness and simplification processes. *Linguistic Inquiry* 26(3): 373–463.

Campbell, Lyle (1974). Phonological features: problems and proposals. *Language* 50(1): 52–65.

Cardona, George, and Babu Suthar (2003). Gujarati. In George Cardona and Dhanesh Jain (eds), *The Indo-Aryan Languages*, 722–65. New York: Routledge.

Chao, Yuen-Ren (1931). Fanqie yu ba zhong [Eight types of Fanqie language]. *Bulletin of the Institute of History and Philology, Academia Sinica* 2(3): 312–54.

Chao, Yuen-Ren (1934). The non-uniqueness of phonemic solutions of phonetic systems. *Bulletin of the Institute of History and Philology, Academia Sinica* 4(4): 363–97.

Chao, Yuen-Ren (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Cherchi, Marcello (1999). *Georgian*. Munich: Lincom Europa.

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

Chumakina, Marina, Greville G. Corbett, and Dunstan Brown (2008). Archi language tutorial (handout). Linguistic Association of Great Britain Meeting, Sept. 2008, University of Essex. http://www.archi.surrey.ac.uk/handout.pdf, accessed Jan. 13, 2014.

Clements, George Nick (1985). The geometry of phonological features. *Phonology Yearbook* 2: 225–52.

Clements, George Nick (1999). Affricates as noncontoured stops. In Osamu Fujimura, Brian D. Joseph, and Bohumil Palek (eds), *Proceedings of LP '98: Item Order in Language and Speech*, vol. 1, 271–99. Prague: Karolinum Press.

Clements, George Nick (2009). The role of features in phonological inventories. In Eric Raimy and Charles E. Cairns (eds), *Contemporary Views on Architecture and Representations in Phonology*, 19–68. Cambridge, Mass.: MIT Press.

Clements, George Nick, and Elizabeth Hume (1995). The internal organization of speech sounds. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, 245–306. Oxford: Blackwell.

Clements, George Nick, and Samuel Jay Keyser (1983). *CV Phonology: A Generative Theory of the Syllable*. Cambridge, Mass.: MIT Press.

Clements, George Nick, and Rachid Ridouane (2011). *Where Do Phonological Features Come From? Cognitive, Physical and Developmental Bases of Distinctive Speech*. Amsterdam: Benjamins.

Cohn, Abigail C. (1993). Nasalisation in English: phonology or phonetics? *Phonology* 10(1): 43–81.

Crewe, W. J. (1975). *The Phonological Features of the Ingessana Language*. Khartoum: University of Khartoum, Institute of African and Asian Studies.

Crosby, K. H. (1944). *An Introduction to the Study of Mende, with 'A Phonetic Introduction to Mende' by Ida C. Ward*. Cambridge: Heffer.

Crothers, John H., James P. Lorentz, Donald Sherman, and Marilyn M. Vihman (1979). *Handbook of Phonological Data from a Sample of the World's Languages: A Report of the Stanford Phonology Archive*. Stanford, Calif.: Dept of Linguistics, Stanford University.

Di Luzio, Aldo (1972). Preliminary description of the Amo language. *Afrika und Ubersee* 56: 3–60.

Dimmendaal, Gerrit Jan (1983). *The Turkana Language*. Cinnaminson, NJ: Foris.

Dimmendaal, Gerrit Jan (1993). Review of *Turkana–English Dictionary* by A. Barrett. *Research in African Literatures* 24(2), special issue on oral literature, 131–3.

Disner, Sandra Ferrari (1983). Vowel quality: the relation between universal and language-specific factors. Doctoral dissertation, UCLA. Distributed as *UCLA Working Papers in Phonetics* 58: 1–158.

Dixon, Robert M. W. (1988). *A Grammar of Boumaa Fijian*. Chicago: University of Chicago Press.

Doke, Clement Martyn, and S. Machabe Mofokeng (1957). *Textbook of Southern Sotho Grammar*. New York: Longmans, Green.

Donohue, Mark (1996). *Warembori*. Munich: Lincom Europa.

Dresher, B. Elan (2009). *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

Duanmu, San (1994). Against contour tone units. *Linguistic Inquiry* 25(4): 555–608.

Duanmu, San (1996). Tone: an overview. *Glot International* 2(4): 3–10.

Duanmu, San (2007). *The Phonology of Standard Chinese*, 2nd edn. Oxford: Oxford University Press.

Duanmu, San (2008). *Syllable Structure: The Limits of Variation*. Oxford: Oxford University Press.

Durvasula, Karthik (2009). Understanding nasality. Doctoral dissertation, University of Delaware.

Elugbe, Ben Ohiomamhe (1989). *Comparative Edoid: Phonology and Lexicon*. Port Harcourt: University of Port Harcourt Press.

Fant, Gunnar (1983). Feature analysis of Swedish vowels: a revisit. *Department for Speech, Music and Hearing Quarterly Progress and Status Report (STL-QPSR)* 24(2–3): 1–19.

Firth, John Rupert (1957). *Papers in Linguistics 1934–1957*. London: Oxford University Press.

Flemming, Edward (1995). Auditory representations in phonology. Doctoral dissertation, University of California, Los Angeles.

Fowler, Carol A. (2012). Segments in articulatory phonology. Talk presented at CUNY Conference on the Segment, CUNY Graduate Center, New York.

Fox, Anthony (1990). *The Structure of German*. Oxford: Clarendon Press.

Fu, Maoji (1956). Beijing hua de yinwei he pinyin zimu [Phonemes and Pinyin symbols in the Peking speech]. *Zhongguo Yuwen* 1956(5): 3–12.

Fudge, Erik C. (1967). The nature of phonological primes. *Journal of Linguistics* 3(1): 1–36.

Fujimura, Osama (1979). An analysis of English syllables as cores and affixes. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 32(4): 471–6.

Gamnes, Hans (1965). En Akustisk undersokelse og maling av formantfrekvenser i ost-norske vokaler. Thesis, University of Oslo.

Garvin, Paul L. (1950). Wichita I: phonemics. *International Journal of American Linguistics* 16(4): 179–84.

Gimson, A. C. (1970). *An Introduction to the Pronunciation of English*, 2nd edn. London: Arnold.

Goldinger, Stephen D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2): 251–79.

Goldsmith, John (1976). Autosegmental phonology. Doctoral dissertation, MIT. Repr. by the Indiana University Linguistics Club, Bloomington.

Goldsmith, John (2011). Syllables. In John A. Goldsmith, Jason Riggle, and Alan C. L. Yu (eds), *Handbook of Phonological Theory*, 164–96. Oxford: Wiley-Blackwell.

Greenberg, Steven, and Eric Zee (1979). On the perception of contour tones. *UCLA Working Papers in Phonetics* 45: 150–64.

Gulya, János (1966). *Eastern Ostyak Chrestomathy*. Bloomington: Indiana University.

Gussmann, Edmund (2002). *Phonology: Analysis and Theory*. Cambridge: Cambridge University Press.

Haiman, John (2011). *Cambodian: Khmer*. Amsterdam: Benjamins.

Hale, Mark, and Charles Reiss (2000). "Substance abuse" and "dysfunctionalism": current trends in phonology. *Linguistic Inquiry* 31: 157–69.

Hall, Daniel Currie (2007). The role and representation of contrast in phonological theory. Doctoral dissertation, University of Toronto.

Halle, Morris (1957). In defense of the number two. In Ernst Pulgram (ed.), *Studies Presented to Joshua Whatmough on his Sixtieth Birthday*, 65–72. The Hague: Mouton.

Halle, Morris (1959). *The Sound Pattern of Russian: A Linguistic and Acoustical Investigation. With an excursus on the contextual variants of the Russian vowels by Lawrence G. Jones*. The Hague: Mouton.

Halle, Morris (1962). Phonology in generative grammar. *Word* 18: 54–72.

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

Halle, Morris (1995). Feature geometry and feature spreading. *Linguistic Inquiry* 26: 1–46.

Halle, Morris (2003). Phonological features. In William J. Frawley (ed.), *International Encyclopedia of Linguistics*, 2nd edn, vol. 3, 314–20. Oxford: Oxford University Press.

Halle, Morris (2005). Palatalization/velar softening: what it is and what it tells us about the nature of language. *Linguistic Inquiry* 36(1): 23–41.

Halle, Morris, and G. Nick Clements (1983). *Problem Book in Phonology*. Cambridge, Mass.: MIT Press.

Halle, Morris, and Kenneth N. Stevens (1969). On the feature "Advanced Tongue Root." *MIT Research Laboratory of Electronics Quarterly Progress Report* 94: 209–15.

Halle, Morris, and Kenneth N. Stevens (1971). A note on laryngeal features. *Quarterly Progress Report* 101: 198–213. Cambridge, Mass.: Research Laboratory of Electronics, MIT.

Halle, Morris, Bert Vaux, and Andrew Wolfe (2000). On feature spreading and the representation of place of articulation. *Linguistic Inquiry* 31(3): 387–444.

Halle, Morris, and Jean-Roger Vergnaud (1987). *An Essay on Stress*. Cambridge, Mass.: MIT Press.

Hamilton, William S. (1980). *Introduction to Russian Phonology and Word Structure*. Columbus, OH: Slavica.

Harris, John (2006). Wide-domain r-effects in English. *UCL Working Papers in Linguistics* 18: 357–79.

Hayes, Bruce (1995). *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.

Hayes, Bruce (1999). Phonetically driven phonology: the role of Optimality Theory and inductive grounding. In Michael Darnell, Edith Moravscik, Michael Noonan, Frederick Newmeyer, and Kathleen Wheatly (eds), *Functionalism and Formalism in Linguistics*, vol. 1: *General Papers*, 243–85. Amsterdam: Benjamins.

Hayes, Bruce (2009). *Introductory Phonology*. Oxford: Wiley-Blackwell.

Heath, Jeffrey (1984). *Functional Grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies. Distributed in North America by Humanities Press, Atlantic Highlands, NJ.

Herbert, Robert K. (1975). Reanalyzing prenasalized consonants. *Studies in African Linguistics* 6(2): 105–23.

Herbert, Robert K. (1986). *Language Universals, Markedness Theory, and Natural Phonetic Processes*. Berlin: Mouton de Gruyter.

Hewitt, B. George (1995). *Georgian: A Structural Reference Grammar*. Amsterdam: Benjamins.

Hewitt, B. George (2006). Caucasian languages. In Keith Brown (ed.), *The Encyclopedia of Language and Linguistics*, 2nd edn, 260–65. Amsterdam: Elsevier.

Hoard, James E. (1971). The new phonological paradigm. *Glossa* 5: 222–68.

Hoffmann, Carl (1963). *A Grammar of the Margi Language*. London: Oxford University Press for the International African Institute.

Hollis, Alfred C. (1905). *The Masai: Their Language and Folklore*. Oxford: Clarendon Press. Republished 1971.

Hombert, Jean-Marie (1978). Consonant types, vowel quality, and tone. In Victoria A. Fromkin (ed.), *Tone: A Linguistic Survey*, 77–111. New York: Academic Press.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan (1979). Phonetic explanation for the development of tones. *Language* 55: 37–58.

House, Arthur S., and Grant Fairbanks (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America* 25(1): 105–13.

Howie, John (1976). *An Acoustic Study of Mandarin Tones and Vowels*. London: Cambridge University Press.

Hualde, Jose Ignacio (1988). Affricates are not contour segments. In Hagit Borer (ed.), *Proceedings of West Coast Conference on Formal Linguistics* 7: 143–57. Stanford, Calif.: Stanford Linguistics Association.

Hualde, Jose Ignacio (1989). Autosegmental and metrical spreading in the vowel-harmony system of northwestern Spain. *Linguistics* 27: 773–805.

Huang, Lijun, and San Duanmu (2013). Xiandai Hanyu ci chang tanxing de lianghua yanjiu [A quantitative study of elastic word length in Modern Chinese]. *Yuyan Kexue* [Linguistic Sciences] 12(1): 8–16.

Hughes, Earl J., and Velma J. Leeding (1971). The phonemes of Nunggubuyu. In *Papers on the Languages of Australian Aboriginals*, 72–81. Canberra: Australian Institute of Aboriginal Studies.

Huisman, Ronald D. (1973). Angaataha verb morphology. *Linguistics* 110(1): 43–54.

Huisman, Ronald D., Roberta Huisman, and Joy Lloyd (1981). Angaatiha syllable patterns. In Phyllis M. Healey (ed.), *Angan Languages Are Different: Four Phonologies*, 51–62. Huntington Beach, Calif.: Summer Institute of Linguistics.

Hyman, Larry (1985). *A Theory of Phonological Weight*. Dordrecht: Foris.

Inkelas, Sharon (1995). The consequences of optimization for underspecification. In Jill Beckman (ed.), *Proceedings of the North East Linguistics Society* 25, 287–302. Amherst, Mass.: GLSA.

Innes, Gordon (1966). *An Introduction to Grebo*. London: School of Oriental and African Studies/Luzac.

Innes, Gorden (1969). *Mende–English Dictionary*. London: Cambridge University Press.

International Phonetic Association (1999). *Handbook of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Ito, Junko, Armin Mester, and Jaye Padgett (1995). Licensing and underspecification in Optimality Theory. *Linguistic Inquiry* 26(4): 571–613.

Iverson, Gregory K. and Joseph C. Salmons (1995). Aspiration and laryngeal representation in Germanic. *Phonology* 12(3): 369–96.

Jakobson, Roman, Gunnar Fant, and Morris Halle (1952). *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge, Mass.: MIT Press.

Jakobson, Roman, and Morris Halle (1964). Tenseness and laxness. In David Abercrombie et al. (eds), *In Honour of Daniel Jones*, 96–101. London: Longman, Green.

Jiang, Zhu-yi (1980). Naxiyu gaikuang [A brief description of the Naxi language]. *Minzu Yuwen* 1980(3): 59–73.

Jones, Daniel (1962). *Outline of English Phonetics*. Cambridge: Heffer.

Jones, Robert B. (1961). *Karen Linguistic Studies*. Berkeley: University of California Press.

Joos, Martin (ed.) (1957). *Readings in Linguistics*. New York: American Council of Learned Societies.

Jørgensen, H.-P. (1969). Die gespannten und ungespannten Vokale in der norddeutschen Hochsprache, mit einer spezifischen Untersuchung der Struktur ihrer Formantenfrequenzen. *Phonetica* 19: 217–45.

Karlgren, Bernhard (1915–26). *Études sur la phonologie chinoise*. Uppsala: Appelberg/Leyden: Brill.

Katamba, Francis (1989). *An Introduction to Phonology*. London: Longman.

Keating, Patricia A. (1985). Universal phonetics and the organization of grammars. In Victoria A. Fromkin (ed.), *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*, 115–32. New York: Academic Press.

Keating, Patricia A. (1988). Underspecification in phonetics. *Phonology* 5: 275–92.

Kehrein, Wolfgang (2006). Complex segments. In Keith Brown (ed.), *Encyclopedia of Language and Linguistics*, 2nd edn, 705–9. Oxford: Elsevier.

Kenstowicz, Michael (1994). *Phonology in Generative Grammar*. Oxford: Blackwell.

Kenstowicz, Michael J. (2009). Two notes on Kinande vowel harmony. *Language Sciences* 31: 248–70.

Kenstowicz, Michael, and Charles Kisseberth (1979). *Generative Phonology*. San Diego, Calif.: Academic Press.

Key, Mary Ritchie (1968). *Comparative Tacanan Phonology*. The Hague: Mouton.

Kim, Mi-Ryoung (2000). Segmental and tonal interactions in English and Korean: a phonetic and phonological study. Doctoral dissertation, University of Michigan, Ann Arbor.

Kim, Mi-Ryoung, Patrice Beddor, and Julie Horrocks (2002). The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30(1): 77–100.

Kim, Mi-Ryoung, and San Duanmu (2004). Tense and lax stops in Korean. *Journal of East Asian Linguistics* 13: 59–104.

Kingston, John (2011). Tonogenesis. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice (eds), *Blackwell Companion to Phonology*, vol. 4, 2304–33. Oxford: Wiley-Blackwell.

Kiparsky, Paul (1974). A note on the vowel features. In Ellen Kaisse and J. Hankamer (eds), *Proceedings of the North-Eastern Linguistic Society V*, 162–71. Cambridge, Mass.: Harvard Linguistics Department.

Klingler, Thomas A. (1992). *A Descriptive Study of the Creole Speech of Pointe Coupee Parish, Louisiana, with Focus on the Lexicon*. Ann Arbor: UMI.

Klingler, Thomas A. (2003). *If I Could Turn My Tongue Like That*. Baton Rouge: Louisiana State University Press.

Kodzasov, Sandro (1996). Phonological preliminaries. In Alexander E. Kibrik (ed.), *Godoberi*, 1–6. Munich: Lincom Europa.

Kohler, Klaus (1966). Towards a phonological theory. *Lingua* 16: 337–51.

Kohler, Klaus (1999). German. In International Phonetic Association, *Handbook of the International Phonetic Association* , 86–9. Cambridge: Cambridge University Press.

Kratochvil, Paul (1970). Book review: *A Grammar of Spoken Chinese* by Y. R. Chao. *Language* 46(2): 513–24.

Kuipers, Aert H. (1960). *Phoneme and Morpheme in Kabardian (Eastern Adyghe)*. The Hague: Mouton.

Kutsch Lojenga, Constance (1994). *Ngiti: A Central-Sudanic Language of Zaire*. Cologne: Köppe.

Ladefoged, Peter (1964). *A Phonetic Study of West African Languages*. Cambridge: Cambridge University Press.

Ladefoged, Peter (1971). *Preliminaries to Linguistic Phonetics*. Chicago: University of Chicago Press.

Ladefoged, Peter (1972). Phonetic prerequisites for a distinctive feature theory. In Albert Valdman (ed.), *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*, 273–86. The Hague: Mouton.

Ladefoged, Peter (1980). What are linguistic sounds made of? *Language* 56(3): 485–502.

Ladefoged, Peter (1992). The many interfaces between phonetics and phonology. In Wolfgang U. Dressler, Hans C. Luschützky, Oskar E. Pfeiffer, and John R. Rennison (eds), *Phonologica 1988*, 165–79. Cambridge: Cambridge University Press.

Ladefoged, Peter (2001). *Vowels and Consonants: An Introduction to the Sounds of Languages*. Malden, Mass.: Blackwell.

Ladefoged, Peter (2007). Articulatory features for describing lexical distinctions. *Language* 83(1): 161–80.

Ladefoged, Peter, Joseph DeClerk, Mona Lindau, and George Papcun (1972). An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics* 22: 48–75.

Ladefoged, Peter, and Morris Halle (1988). Some major features of the International Phonetic Alphabet. *Language* 64(3): 577–82.

Ladefoged, Peter, and Keith Johnson (2011). *A Course in Phonetics*, 6th edn. Independence, Ky.: Cengage.

Ladefoged, Peter, and Ian Maddieson (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Lass, Roger (1984). Vowel system universals and typology: prologue to theory. *Phonology Yearbook* 1: 75–112.

Lawrence, Wayne P. (2000). /str/ → /ʃtr/: assimilation at a distance? *American Speech* 75.1: 82–7.

Leben, William (1973). Suprasegmental phonology. Doctoral dissertation, MIT.

Lee, Wai-Sum, and Eric Zee (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association* 33(1): 109–12.

Lehiste, Ilse (1960). Segmental and syllabic quantity in Estonian. *American Studies in Uralic Linguistics* 1: 21–82.

Levin, Juliette (1985). A metrical theory of syllabicity. Doctoral dissertation, MIT.

Levine, Robert V., and Ara Norenzayan (1999). The pace of life in 31 countries. *Journal of Cross-Cultural Psychology* 30(2): 178–205.

Lewis, Geoffrey L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.

Li, Fang-Kuei (1977). *A Handbook of Comparative Tai*. Honolulu: University of Hawaii Press.

Liberman, Alvin M., Franklin S. Cooper, Donald P. Shankweiler, and Michael Studdert-Kennedy (1967). Perception of the speech code. *Psychological Review* 74(6): 431–61.

Liberman, Alvin M., and Ignatius G. Mattingly (1985). The motor theory of speech perception revised. *Cognition* 21: 1–36.

Liberman, Mark, and Janet Pierrehumbert (1984). Intonational invariance under changes in pitch range and length. In Mark Aronoff, Richard Oehrle, Frances Kelley, and Bonnie Wilker Stephens (eds), *Language Sound Structure: Studies in Phonology Presented to Morris Halle by his Teacher and Students*, 157–233. Cambridge, Mass.: MIT Press.

Lin, Maocan, and Jingzhu Yan (1988). The characteristic features of the final reduction in the neutral-tone syllable of Beijing Mandarin. *Phonetic Laboratory Annual Report of Phonetic Research*, 37–51. Beijing: Phonetic Laboratory, Institute of Linguistics, Chinese Academy of Social Sciences.

Lin, Yen-Hwei (2011). Affricates. In Marc Van Oostendorp, Colin Ewen, Elizabeth Hume, and Keren Rice (eds), *The Blackwell Companion to Phonology*, vol. 1, 367–90. Oxford: Wiley-Blackwell.

Lindau, Mona (1979). The feature expanded. *Journal of Phonetics* 7: 163–76.

Lombardi, Linda (1990). The nonlinear organization of the affricate. *Natural Language and Linguistic Theory* 8: 374–425.

Lowin, Aaron, Joseph H. Hottes, Bruce E. Sandler, and Marc Bornstein (1971). The pace of life and sensitivity to time in urban and rural settings: a preliminary study. *Journal of Social Psychology* 83(2): 247–53.

Maddieson, Ian (1984). *Patterns of Sounds*. London: Cambridge University Press.

Maddieson, Ian (1987). Revision of the IPA: linguo-labials as a test case. *Journal of the International Phonetic Association* 17: 26–30.

Maddieson, Ian (1989). Prenasalized stops and speech timing. *Journal of the International Phonetic Association* 19(2): 57–66.

Maddieson, Ian (2005). Vowel quality inventories. In Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie (eds), *The World Atlas of Language Structures*, 14–15. Oxford: Oxford University Press.

Maddieson, Ian, and Kristin Precoda (1990). Updating UPSID. *UCLA Working Papers in Phonetics* 74: 104–11.

Malécot, André (1960). Vowel nasality as a distinctive feature in American English. *Language* 36(2): 222–9.

Marconnès, Revd Francisque (1931). *A Grammar of Central Karanga*. Johannesburg: Witwatersrand University Press.

Mawadza, Aquilina M. (2009). *A Basic Ndebele Grammar*. Hyattsville, Md.: Dunwoody Press.

McCarthy, John (1979). Formal problems in Semitic phonology and morphology. Doctoral dissertation, MIT.

McClean, Reginald John (1947). *Swedish: A Grammar of the Modern Language*. London: Hodder and Stoughton.

McMahon, April (2002). *An Introduction to English Phonology*. Oxford: Oxford University Press.

Mielke, Jeff (2004–7). P-base v1.92. http://aix1.uottawa.ca/~jmielke/pbase/

Mielke, Jeff (2008). *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

Miller, Amanda (2011). The representation of clicks. In Marc Van Oostendorp, Colin J. Ewen, Elizabeth Hume, and Keren Rice, *The Blackwell Companion to Phonology*, vol. 1, 416–39. Oxford: Wiley-Blackwell.

Mohrlang, Roger (1972). *Higi Phonology*. Zaria: Institute of Linguistics.

Montgomery, C. (1970). Problems in the development of an orthography for the Sebei language of Uganda. *Journal of the Language Association of East Africa* 1: 48–55.

Moseley, Christopher (ed.) (2010). *Atlas of the World's Languages in Danger*, 3rd edn. Paris: UNESCO. http://www.unesco.org/culture/en/endangeredlanguages/atlas, accessed Jan. 12, 2014.

Moulton, William G. (1956). Syllabic nuclei and final consonant clusters in German. In Morris Halle, Horace G. Lunt, Hugh McLean, and Cornelis H. van Schooneveld (eds), *For Roman Jakobson: Essays on the Occasion of his Sixtieth Birthday*, 372–81. The Hague: Mouton.

Newmark, Leonard (1957). *Structural Grammar of Albanian*. Bloomington: Indiana University Research Center in Anthropology, Folklore and Linguistics.

Ngo, Binh Nhu (1999). *Elementary Vietnamese*. North Clarendon, Vt: Tuttle.

Ngo, Binh Nhu (2001). The Vietnamese language learning framework, pt 1: Linguistics. *Journal of Southeast Asian Language Teaching* 10: 1–24. http://www.seasite.niu.edu/jsealt/, accessed May 31, 2012.

Nielsen, Kuniko Yasu (2008). The specificity of allophonic variability and its implications for accounts of speech perception. Doctoral dissertation, UCLA.

Nielsen, Kuniko Yasu (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39(2): 132–42.

Nihalani, Paroo (1986). Phonetic implementation of implosives. *Language and Speech* 29(3): 253–62.

Nordbustad, Frøydis (1988). *Iraqw Grammar: An Analytical Study of the Iraqw*. Berlin: Reimer.

Ohala, John J. (1975). Phonetic explanations for nasal sound patterns. In Charles A. Ferguson, Larry M. Hyman, and John J. Ohala (eds), *Nasalfest: Papers from a Symposium on Nasals and Nasalization*, 289–316. Stanford, Calif.: Language Universal Project.

Ohala, John J., and Manjari Ohala (1993). The phonetics of nasal phonology: theorems and data. In Marie K. Huffman and Rena Krakow (eds), *Phonetics and Phonology 5: Nasals, Nasalization, and the Velum*, 225–49. San Diego, Calif.: Academic Press.

Olson, Kenneth S., Duane Reiman, Fernando Sabio, and Felipe da Silva (2009). The voiced linguolabial plosive in Kajoko. *Proceedings from the 45th Annual Meeting of the Chicago Linguistic Society*, 519–30.

Padgett, Jaye (1995). *Stricture in Feature Geometry*. Stanford, Calif.: Center for the Study of Language and Information, Stanford University.

Parkinson, Frederick B. (1996). The representation of vowel height in phonology. Doctoral dissertation, Ohio State University, Columbus.

Pelling, J. N. (1975). *A Practical Ndebele Dictionary*, 2nd impression. Salisbury: Longman Rhodesia.

Perkell, Joseph S. (1969). *Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study*. Cambridge, Mass.: MIT Press.

Perkell, Joseph S. (1971). Physiology of speech production: a preliminary study of two suggested revisions of the features specifying vowels. *Quarterly Progress Report* 102: 123–39. Cambridge, Mass. : Research Laboratory of Electronics, MIT.

Peterson, Gordon E., and Harold L. Barney (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America* 24(2): 175–84.

Peterson, Gordon E., and Ilse Lehiste (1960). Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* 32(6): 693–703.

Pitt, Mark A., Laura Dilley, Keith Johnson, Scott Kiesling, William Raymond, Elizabeth Hume, and Eric Fosler-Lussier (2007). *Buckeye Corpus of Conversational Speech* (2nd release) [www.buckeyecorpus.osu.edu, accessed Jan. 14, 2014] Columbus: Dept of Psychology, Ohio State University.

Polley, Linda, and Dorothea Jeffrey (1977). *Phonological Structure of Turkana: A Report on Some Aspects of Vowel Harmony and Tone*. Nairobi: Summer Institute of Linguistics.

Port, Robert F., and Adam P. Leary (2005). Against formal phonology. *Language* 81(4): 927–64.

Postal, Paul M. (1968). *Aspects of Phonological Theory*. New York: Harper & Row.

Press, Ian (1987). *A Grammar of Modern Breton*. New York: Mouton de Gruyter.

Priest, Perry N. (1968). Phonemes of the Siriono language. *Linguistics* 41:102–8.

Prince, Alan (1980). A metrical theory for Estonian quantity. *Linguistic Inquiry* 11: 511–62.

Prince, Alan (1990). Quantitative consequences of rhythmic organization. In *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, vol. 2: *The Parasession on the Syllable in Phonetics and Phonology*, 355–98. Chicago: Chicago Linguistic Society.

Pulleyblank, Douglas (1983). Tone in lexical phonology. Doctoral dissertation, MIT.

Pulleyblank, Edwin G. (1997). The Cantonese vowel system in historical perspective. In Jialing Wang and Norval Smith (eds), *Studies in Chinese Phonology*, 185–218. Berlin: Mouton de Gruyter.

Ray, Punya Sloka (1967). Dafla phonology and morphology. *Anthropological Linguistics* 9(8): 9–14.

Ray, Punya Sloka, Muhammad Abdul Hai, and Lila Ray (1966). *Bengali Language Handbook*. Washington, DC: Center for Applied Linguistics.

Redden, James E. (1979). *A Descriptive Grammar of Ewondo*. Department of Linguistics, South Illinois University, Carbondale.

Rice, Keren (1992). On deriving sonority: a structural account of sonority relationships. *Phonology* 9(1): 61–99.

Rice, Keren D. (1995). On vowel place features. *Toronto Working Papers in Linguistics* 14(1): 73–116.

Rice, Keren (2003). Review of T. Alan Hall (ed.), *Distinctive Feature Theory*. *Journal of Linguistics* 39: 400–405.

Rice, Keren (2006). Markedness in phonology. In Paul de Lacy (ed.), *The Cambridge Handbook of Phonology*, 79–97. Cambridge: Cambridge University Press.

Robins, R. H., and Natalie Waterson (1952). Notes on the phonetics of the Georgian word. *Bulletin of the School of Oriental and African Studies* 15: 55–72.

Rood, David S. (1975). The implications of Wichita phonology. *Language* 51(2): 315–37.

Rood, David S. (1976). *Wichita Grammar*. New York: Garland.

Ruhlen, Merritt (1976). *A Guide to the Languages of the World*. Stanford, Calif.: Stanford University.

Sagey, Elizabeth (1986). The representation of features and relations in nonlinear phonology. Doctoral dissertation, MIT.

Samely, Ursula (1991). *Kedang (Eastern Indonesia): Some Aspects of its Grammar*. Hamburg: Buske.

Sancier, Michele L., and Carol A. Fowler (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25(4): 421–36.

Santos, Rosine (1977). *Phonologie et morphotonologie de la langue w̃eẽỹ (koñagi)*. Dakar: Centre de linguistique appliquée de Dakar.

Sapir, Edward (1931). Notes on the Gweabo language of Liberia. *Language* 7(1): 30–41.

Schuh, Russell G. (1972). Aspects of Ngizim syntax. University of California, Los Angeles.

Selkirk, Elisabeth (1982). The syllable. In Harry van der Hulst and Norval Smith (eds), *The Structure of Phonological Representations*, pt 2, 337–83. Dordrecht: Foris.

Simpson, Adrian P. (1999). Fundamental problems in comparative phonetics and phonology: does UPSID help to solve them? In *Proceedings of ICPhS99* (San Francisco), 349–52.

Singler, John Victor (1979). The segmental phonology of verb suffixes in Talo Klao (Kru). MA thesis, University of California, Los Angeles.

Singler, John Victor (1983). Vowel harmony in Klao: linear and nonlinear analyses. *Studies in African Linguistics* 14: 1–33.

Spears, Richard A. (1967). Tone in Mende. *Journal of African Languages* 6: 231–44.

Spotts, Hazel (1953). Vowel harmony and consonant sequences in Mazahua1 (Otomi). *International Journal of American Linguistics* 19(4): 253–8.

Stanley, Richard (1967). Redundancy rules in phonology. *Language* 43(2): 393–436.

Steriade, Donca (1987). Redundant values. In *Papers from the 23rd Annual Regional Meeting of the Chicago Linguistic Society*, pt 2: *Parasession on Autosegmental and Metrical Phonology*, 339–62. Chicago: CLS.

Steriade, Donca (1989). Affricates are stops. Paper presented at Conference on Features and Underspecification Theories, Oct. 7–9, MIT.

Steriade, Donca (1993). Closure, release, and nasal contours. In Marie K. Huffman and Rena A. Krakow (eds), *Nasals, Nasalization, and the Velum*, 401–70. San Diego, Calif.: Academic Press.

Steriade, Donca (1995). Underspecification and markedness. In John Goldsmith (ed.), *The Handbook of Phonological Theory*, 114–74. Oxford: Blackwell.

Steriade, Donca (1999). Phonetics in phonology: the case of laryngeal neutralization. In Matthew Gordon (ed.), *Phonetically Driven Phonology*, 25–145. Los Angeles: Department of Linguistics, University of California.

Stevens, Kenneth (1972). The quantal nature of speech: evidence from articulatory-acoustic data. In Edward E. David and Peter B. Denes (eds), *Human Communication: A Unified View*, 51–66. New York: McGraw-Hill.

Stevens, Kenneth (1989). On the quantal nature of speech. *Journal of Phonetics* 17(1): 3–45.

Stewart, J. M. (1967). Tongue root position in Akan vowel harmony. *Phonetica* 16: 185–204.

Stirtz, Timothy M. (2012). A grammar of Gaahmg, a Nilo-Saharan language of Sudan. Doctoral dissertation, Leiden University.

Stokhof, W. A. L. (1979). *Woisika II: Phonemics*. Canberra: Australian National University.

Stone, Maureen, and Andrew Lundberg (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of the Acoustical Society of America* 99(6): 3728–37.

Stringer, Mary, and Joyce Hotz (1973). Waffa phonemes. In Howard McKaughan (ed.), *The Languages of the Eastern Family of the East New Guinea Highland Stock*, 523–9. Seattle: University of Washington Press.

Thompson, Laurence C. (1965). *A Vietnamese Grammar*. Seattle: University of Washington Press.

Thorne, David A. (1993). *A Comprehensive Welsh Grammar*. Oxford: Blackwell.

Traill, Anthony (1985). *Phonetic and Phonological Studies of !XÓÕ Bushman*. Hamburg: Buske.

Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie*. Prague: Le Cercle.

Trubetzkoy, Nikolai Sergeevich (1969). *Principles of Phonology*. Translated from German by Christiane A. M. Baltaxe. Berkeley: University of California Press.

Tryon, Darrell T. (1976). *New Hebrides Languages: An Internal Classification*. Canberra: Dept of Linguistics, Research School of Pacific Studies, Australian National University.

Tucker, Archibald N. (1962). The syllable in Luganda: a prosodic approach. *Journal of African Languages* 1: 122–66.

Tucker, Archibald N. (1971). Notes on Ik. *African Studies* 30(3–4): 341–54.

Tucker, Archibald N., and M. A. Bryan (and James Woodburn as co-author for Hadza) (1977). The East African click languages: a phonetic comparison. In J. G. Moehlig, Franz Rottland, and Bernd Heine (eds), *Zur Sprachgeschichte und Ethnohistorie in Afrika*, 300–323. Berlin: Diener.

Tucker, Archibald N., and J. Tompo Ole Mpaayei (1955). *A Maasai Grammar with Vocabulary*. London: Longmans, Green.

Unbegaun, Boris Ottokar (1957). *Russian Grammar*. Oxford: Clarendon Press.

Valdman, Albert, Thomas A. Klingler, Margaret Marshall, and Kevin J. Rottet (eds) (1998). *Dictionary of Louisiana Creole*. Bloomington: Indiana University Press.

van der Hulst, Harry, and Norval Smith (1982). An overview of autosegmental and metrical phonology. In Harry van der Hulst and Norval Smith (eds), *The Structure of Phonological Representations*, pt 1, 1–45. Dordrecht: Foris.

Van Wynen, Donald, and Mabel Garrard de Van Wynen (1962). *Fonemas tacana y modelos de acentuacion*. Cochabamba: Summer Institute of Linguistics.

Vaux, Bert (1996). The status of ATR in feature geometry. *Linguistic Inquiry* 27(1): 175–82.

Vaux, Bert (2009). The role of features in a symbolic theory of phonology. In Eric Raimy and Charles E. Cairns (eds), *Contemporary Views on Architecture and Representations in Phonology*, 75–97. Cambridge, Mass.: MIT Press.

Verheijen, J. A. J. (1986). *The Sama/Bajau Language in the Lesser Sunda Islands*. Canberra: Dept of Linguistics, Research School of Pacific Studies, Australian National University.

Walker, Rachel (1993). A vowel feature hierarchy for contrastive specification. In Carrie Dyck (ed.), *Toronto Working Papers in Linguistics* 12(2): 178–97.

Walker, Rachel, and Geoffrey K. Pullum (1999). Possible and impossible segments. *Language* 75(3): 764–80.

Wang, William S.-Y. (1967). Phonological features of tone. *International Journal of American Linguistics* 33(2): 93–105.

Ward, Ida C. (1933). *The Phonetic and Tonal Structure of Efik*. Cambridge: Heffer.

Ward, Ida C. (1944). A phonetic introduction to Mende. In K. H. Crosby, *An Introduction to the Study of Mende*, 1–7. Cambridge: Heffer.

Watson, Richard L. (1964). Pacoh phonemes. *Mon-Khmer Studies* 1: 135–48.

Watson, Richard L. (1996). Why three phonologies for Pacoh? *Mon-Khmer Studies* 26: 197–205.

Werle, Johannes Martin, and Dagou Justin Gbalehi (1976). *Phonologie et morphonologie du Bété de la région de Guiberoua*. Abidjan: Institut de linguistique appliquée, Université d'Abidjan.

Wichmann, Søren (1995). *The Relationship Among the Mixe-Zoquean Languages of Mexico*. Salt Lake City: University of Utah Press.

Wiese, Richard (1996). *The Phonology of German*. Oxford: Oxford University Press.

Wiesemann, Ursula (1972). *Die phonologische und grammatische Struktur der Kaingang-Sprache*. The Hague: Mouton.

Williams, Edwin (1976). Underlying tone in Margi and Igbo. *Linguistic Inquiry* 7(3): 436–68.

Williamson, Kay (1972). Assimilation in Ogbia. *Research Notes* 5(2–3): 1–5. Ibadan: Dept of Linguistics and Nigerian Languages, University of Ibadan.

Woo, Nancy (1969). Prosody and phonology. Doctoral dissertation, MIT.

Wood, Sidney A. J. (1982). X-ray and model studies of vowel articulation. Doctoral dissertation, Dept of Linguistics, Lund University.

Xu, Lin, and Yansun Zhao (eds) (1984). *Baiyu jianzhi* [A short description of Bai]. Beijing: Minzu Chubanshe.

Yadav, Ramawatar (1996). *A Reference Grammar of Maithili*. Berlin: Mouton de Gruyter.

Yang, Tongyin (2009). *Tongdao Dongyu yanjiu: gongneng shiye xia de yuyin, yufa, he yupian yanjiu* [Aspects of the Kam language, as revealed in its narrative discourse]. Beijing: Minzu Chubanshe.

Yip, Moira (1980). Tonal phonology of Chinese. Doctoral dissertation, MIT.

Yip, Moira (1982). Reduplication and C–V skeleta in Chinese secret languages. *Linguistic Inquiry* 13: 637–61.

Yip, Moira (1989). Contour tones. *Phonology* 6: 149–74.

Yip, Moira (2002). *Tone*. Cambridge: Cambridge University Press.

You, Rujie, Nairong Qian, and Zhengxia Gao (1980). Lun Putonghua de yinwei xitong [On the phonemic system of Standard Chinese]. *Zhongguo Yuwen* 5(158): 328–34.

Zemlin, Willard R. (1998). *Speech and Hearing Science: Anatomy and Physiology*, 4th edn. Boston, Mass.: Allyn and Bacon.

Ziervogel, Dirk (1959). *A Grammar of Northern Transvaal Ndebele*. Pretoria: van Schaik.

Zimmer, Karl, and Orhan Orgun (1992). Turkish. *Journal of the International Phonetic Association* 22(1–2): 43–5.

Author Index

Language Index

Subject Index