Understanding Phonology

Carlos Gussenhoven and Haike Jacobs

Understanding Language Series

Fourth Edition

Routledge
Understanding Phonology

*Understanding Phonology, Fourth Edition* provides a clear, accessible and broad introduction to Phonology. Introducing basic concepts, it provides a comprehensive account of phonological issues such as segmental contrasts; syllables and moras; quantity, tone, intonation and stress; feature geometry; and prosodic constituent structure. This new edition has been reorganized and revised with key features including:

- A brand new eResource at www.routledge.com/9781138961425, which contains a full answer key for all exercises, and audio recordings of illustrative examples;
- Illustrations in languages from all six continents and all major language families, including Arabic, Mandarin, Finnish, Zulu and Hawaiian;
- Over 140 exercises to test understanding, including new exercises involving larger data sets;
- Revised coverage of tone, stress and opacity in OT.

*Understanding Phonology* is essential reading for students coming to this topic for the first time.

**Carlos Gussenhoven** is emeritus professor of general and experimental phonology at Radboud University Nijmegen, Netherlands.

**Haike Jacobs** is professor of French linguistics at Radboud University Nijmegen, Netherlands.
Understanding Language Series

Series Editors:

Bernard Comrie, University of California, Santa Barbara, USA
Greville Corbett, Surrey Morphology Group, University of Surrey, UK

The Understanding Language series provides approachable, yet authoritative, introductions to major topics in linguistics. Ideal for students with little or no prior knowledge of linguistics, each book carefully explains the basics, emphasizing understanding of the essential notions rather than arguing for a particular theoretical position.

Other titles in the series:

Understanding Pragmatics
Gunter Senft

Understanding Child Language Acquisition
Caroline Rowland

Understanding Semantics, Second Edition
Sebastian Löhner

Understanding Syntax, Fourth Edition
Maggie Tallerman

Understanding Phonetics
Patricia Ashby

Understanding Morphology, Second Edition
Martin Haspelmath and Andrea D. Sims

Understanding Language Testing
Dan Douglas

Understanding Second Language Acquisition
Lourdes Ortega

Study Skills for Linguistics
Jeanette Sakel

Understanding Language Change
Kate Burridge and Alexander Bergs

Understanding Phonology, Fourth Edition
Carlos Gussenhoven and Haike Jacobs

For more information on any of these titles, or to order, go to www.routledge.com/series/ULAN
Contents

Preface vi
Acknowledgements viii
The IPA chart ix

1 Structures in languages 1
2 The production of speech 17
3 Some typology: sameness and difference 36
4 Making the form fit 48
5 Distinctive features 64
6 Describing places of articulation 76
7 Connecting underlying and surface representations 94
8 Levels of representation 109
9 Between the segment and the syllable 126
10 Tones 148
11 Word stress 174
12 Phonology above the word 200
13 Transparency and opacity with rules and constraints 215
14 Feature geometry 232

References 259
Language index 273
Subject index 281
There are about 7,000 languages in the world today. Almost certainly, no two of them have the same sound structure: they vary widely in the number of consonants and vowels they have, in their use of tonal contrasts, in their stress patterns, in the shape of their syllables, and so on. At the same time, all these languages show striking similarities in the way they structure their sound systems. Phonology is a thriving field of linguistic research that strives to understand the structure behind these systems. How do these similarities arise? And why, at the same time, is there so much variation? How is our knowledge of the pronunciation of our language represented in our brain? How can we describe the pronunciation of a language? What do people do when they play language games? Why do loanwords often sound so totally different from the way they are pronounced in the donor languages? These and many other questions are dealt with in this book. In our discussion, we have tried to sketch the development of scientific thinking about the sound structure of languages and to take an unbiased view of the cognitive or physiological nature of the explanations. We hope we have succeeded in this task in at least some places in the book, and have got close enough to this ideal for it to serve as a reliable and relevant introduction to an important and exciting field.

This book reflects the thoughts and discoveries of many phonologists. We have learnt to appreciate the value and implications of these theoretical positions not only by reading their publications but also by attending their classes and discussing the issues with them. Needless to say, our debt to them is inestimable.


The second edition (2005) of this book benefited from the comments made by Elan Dresher, San Duanmu, Ed Flemming, Bo Hagström, Victoria Rosén, Ingmar Steiner, Rik van Gijn, Jeroen van de Weijer, Leo Wetzels, Maria Wolters and those who responded to the publisher’s questionnaire.

In the third edition (2011), we made a number of additions and in part rearranged the old text. A new chapter 1 was added to place our subject in a more general scientific context. It was expanded with the first half of the old chapter 2 on the difference between morphosyntactic structure and phonological structure. Chapter 2 is essentially the old chapter 1, while chapter 3 is a revised version of the old chapter 2. Chapter 8 was added in order to deal with responses to phonological
opacity in Optimality Theory. The authors would like to thank Janine Berns, Hyong Sil Cho, Bert Cranen, Anne Cutler, Marinda Hagen, Robert Kennedy, Hikaru Osawa and Henning Reetz for their help with this third edition.

The main conceptual change in this fourth edition (2017) is a more consistent autosegmental treatment of three classes of segments (vowels, consonants and tones) and of their inclusion in the prosodic structure. This has led to a rearrangement of the chapters on syllables, tone and stress, as well as a rewriting of the last two of these chapters. While the functional roles of tones are emphasized throughout the discussion, as a class of segments tones are treated in a single chapter on lexical and intonational tones. The chapter on stress has been expanded and deals with typological issues as much as with theoretical accounts. These rewritings have benefited from the review by David Deterding in JIPA 43 (2013). In addition, we have split the chapter on distinctive features into two, so that students will experience a more even workload across chapters. The new chapter 5 deals with the major-class, laryngeal and manner features, while chapter 6 treats place features and includes the rule format conventions that in earlier editions were discussed in a chapter on Dutch diminutives. That chapter has been discarded in this edition. The chapter on underlying and surface representations is now immediately followed by the chapter on Lexical Phonology. Finally, we have collapsed the two chapters on feature geometry into one, and reduced the discussion on opacity. The chapters on opacity and feature geometry are the last chapters in the book. As a result of these changes, the number of chapters has been reduced from 16 to 14. We thank Mirjam Broersma, Hyong Sil Cho, Bernard Comrie, Grev Corbett, Catia Cucchiariini, Gunnar Holmstedt, Judith Hanssen, Beste Kamali, Hamed Rahmani, Makiko Sadakata and Leo Wetzels for comments on a draft version and help with data. Finally, it was an unspeakably reassuring experience to find Helen Tredget and Katherine Wetzel skilfully seeing our manuscript through production.

Carlos Gussenhoven and Haike Jacobs
Radboud University Nijmegen, the Netherlands
July 2016
Acknowledgements

The International Phonetic Alphabet (2015) is reproduced by kind permission of the International Phonetic Association.
# THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

**CONSONANTS (PULMONIC)** © 2015 IPA

<table>
<thead>
<tr>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Postalveolar</th>
<th>Retrolabial</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>t̜</td>
<td>q̠</td>
<td>k̠</td>
<td>q̡</td>
<td>q̢</td>
<td>?</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>ñ</td>
<td>n̬</td>
<td>n̬̂</td>
<td>ñ̬</td>
<td>n̬̄</td>
<td>n̬̅</td>
<td>n̬̆</td>
<td>ṅ̬</td>
<td>n̬</td>
</tr>
<tr>
<td>Trill</td>
<td>b</td>
<td>r̠</td>
<td>r̡</td>
<td>r̢</td>
<td>ṛ</td>
<td>r̫</td>
<td>r̬</td>
<td>r̭</td>
<td>r̮</td>
<td>r̯</td>
</tr>
<tr>
<td>Tap or Flap</td>
<td>v̠</td>
<td>f̠</td>
<td>θ̠</td>
<td>ð̠</td>
<td>s̠</td>
<td>z̠</td>
<td>j̠</td>
<td>α̠</td>
<td>θ̡</td>
<td>θ̢</td>
</tr>
<tr>
<td>Lateral fricative</td>
<td>ɭ̠</td>
<td>ʃ̠</td>
<td>s̠̩</td>
<td>z̠̩</td>
<td>s̠̫</td>
<td>z̠̬</td>
<td>j̠̩</td>
<td>i̠̩</td>
<td>j̠̫</td>
<td>j̠̬</td>
</tr>
<tr>
<td>Lateral approximant</td>
<td>v̠</td>
<td>ɾ̠</td>
<td>θ̠</td>
<td>ð̠</td>
<td>s̠</td>
<td>z̠</td>
<td>j̠</td>
<td>α̠</td>
<td>θ̡</td>
<td>θ̢</td>
</tr>
<tr>
<td>Approximant</td>
<td>l̠</td>
<td>ɾ̠</td>
<td>θ̠</td>
<td>ð̠</td>
<td>s̠</td>
<td>z̠</td>
<td>j̠</td>
<td>α̠</td>
<td>θ̡</td>
<td>θ̢</td>
</tr>
</tbody>
</table>

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

**CONSONANTS (NON-PULMONIC)**

<table>
<thead>
<tr>
<th>Clicks</th>
<th>Voiced implosives</th>
<th>Ejectives</th>
<th>VOWELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilabial</td>
<td>Blah</td>
<td>Blah</td>
<td>Examples:</td>
</tr>
<tr>
<td>Dental</td>
<td>Dental/alveolar</td>
<td>Dental/alveolar</td>
<td>p̠</td>
</tr>
<tr>
<td>Palatal</td>
<td>Velar</td>
<td>Velar</td>
<td>k̠</td>
</tr>
<tr>
<td>Alveolar lateral</td>
<td>Uvular</td>
<td>Alveolar fricative</td>
<td>s̠</td>
</tr>
</tbody>
</table>

**OTHER SYMBOLS**

- Voiced labio-velar fricative: cz
- Alveolo-palatal fricatives: ʃ, ʒ

**DEACRITICS** Some diacritics may be placed above a symbol with a descender, e.g. ʃ̠.

**TONES AND WORD ACCENTS**

<table>
<thead>
<tr>
<th>LEVEL</th>
<th>CONTOUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>•</td>
<td>Extra high</td>
</tr>
<tr>
<td>•</td>
<td>High</td>
</tr>
<tr>
<td>•</td>
<td>Falling</td>
</tr>
<tr>
<td>•</td>
<td>Mid</td>
</tr>
<tr>
<td>•</td>
<td>Low</td>
</tr>
<tr>
<td>•</td>
<td>Rising</td>
</tr>
<tr>
<td>•</td>
<td>Falling</td>
</tr>
</tbody>
</table>

| • | Downstep |
| • | Upstep |

| • | ʃ̠ |
| • | p̠ |
| • | t̜ |
| • | q̠ |
| • | k̠ |
| • | q̢ |
| • | ? |
| • | m̃ |
| • | n̬̂ |
| • | ñ̬ |
| • | n̬̄ |
| • | n̬̅ |
| • | n̬̆ |
| • | ṅ̬ |
| • | b̃ |
| • | r̠ |
| • | r̡ |
| • | r̢ |
| • | ṛ |
| • | r̫ |
| • | r̬ |
| • | r̭ |
| • | r̮ |
| • | r̯ |
| • | v̠ |
| • | f̠ |
| • | θ̠ |
| • | ð̠ |
| • | s̠ |
| • | z̠ |
| • | j̠ |
| • | α̠ |
| • | θ̡ |
| • | θ̢ |

Where symbols appear in pairs, the one to the right represents a rounded vowel.

**SUPRASEGMENTALS**

- Primary stress |
- Secondary stress |
- Long |
- Half-long |
- Extra-short |
- Minor (foot) group |
- Major (intonation) group |
- Syllable break |
- Linking (absence of a break)
1.1 INTRODUCTION

Imagine a biologist wandering in a newly discovered forest inhabited by thousands of undescribed animal species. Any research he or she takes on will yield completely new scientific data. This privileged situation is very much that of today’s phonologist. At least half of the 7,000 or so languages in the world either have sketchy accounts devoted to them or are entirely undescribed. No two languages have ever been found that have the same phonology. In fact, varieties of the same language often differ in their phonological systems, and there are therefore numerous opportunities to be the first in history to know about some phonological phenomenon. Our hope is that you will be able to derive some excitement even from the existing language data that are presented in this book and that you can see how the ways humans organize their vocal resources into grammars are quite remarkable. Admittedly, seeing those structures, whether in new or existing data, requires more than a quick look. Our general aim in this book is to guide you through the various aspects of phonological structure and indicate how they vary across languages.

Let’s begin by observing that all human languages have two co-existing structures: a phonological structure and a morphosyntactic structure. Before this point can be made, we need to make it clear what it means for languages to have ‘structure’ to begin with. In section 1.2, we discuss how languages vary in the extent to which they allow particular kinds of structure to be ‘seen’, or observed. Phonological structure is not the same as the orthography in alphabetic writing systems, and we urge you to keep the notions of ‘letter’ and ‘sound’ distinct in your thinking about pronunciation. In section 1.4, we explain what is meant by morphosyntactic structure and then move on to a thought experiment in which you are invited to imagine a world without phonological structure, a mental exercise that is intended to make you see more clearly what it is. Its independence from the morphosyntactic structure is brought out by another thought experiment, in which we imagine a world where all languages have the same phonological structure. Finally, we will make the point that all languages have phonological structures, but that sign languages express their phonological elements visually, as manual and facial gestures, rather than acoustically.

1.2 OBSERVING LINGUISTIC STRUCTURE

By the time he or she is five years old, every child in this world has learnt to speak a human language. Today, this event happens about 16,000 times every hour.¹ It is hard to say how many languages spoken in the world today are still being used by
care-givers to communicate with their children, but one thing is certain: those languages vary greatly in their structure. Some of them will have a passive verb form, and others will not; some will have 5 vowels, some 13 and yet others 25, and likewise the number of consonants will vary greatly; some will use pitch to distinguish words (see section 2.2.4), and some will not, and so on. Children of that age are usually capable of saying that they have five fingers on each of their two hands, but if you were to ask them how many vowels their language has, your question would be royally ignored. This is not because the child does not necessarily know the meaning of the word 'vowel', but because the question goes well beyond what humans can naturally know. The structure of our language is not accessible to us in the way the outward shape of our body is; in fact, people are normally not even aware that their language has any structure at all.

1.2.1 Awareness of language structure

How, then, can we ever develop an awareness of that structure? In countries in which children are taught to write in an alphabetic orthography, awareness of segments typically arises because to a large extent the letters used to write words stand in some regular relationship to the sounds – we will often call them 'segments' – in those words. It is in fact quite a mental step to realize that an English word like tea consists of two segments, a [t] and an [ː:], rather than being a single unit of sound. (The symbol [:], known as the length mark, indicates that the preceding sound is long; see chapter 2.) In general, awareness of structural elements will depend on two factors. First, besides your natural inclination to look into such matters, there are the demands that are made on you to do so, as will happen at school. Illiterates may be unaware of the existence of segments, and it will take more than a little work to reach that understanding (Morais et al. 1979). Second, the language itself will reveal some elements of structure more readily than others. That is, the structural elements of any one language vary in 'salience', and it is understandably easier to become aware of more striking elements than of less striking elements. For instance, the notion 'lexical word' naturally develops fairly easily for speakers of most languages. Pre-school speakers of English usually know that Johnny shouted! consists of two words, and No! of one, even though they will be less sure in cases like Don’t!, shell shocked or shaggy-dog story.

1.2.2 Language diversity

But again, we must remember that languages vary, and thus also vary in the salience of comparable notions. Inuit, spoken in northern Alaska, northern Canada and western Greenland, has a complex system of suffixes and incorporates nouns with verbs: many of its words or sentences are much like Don’t!, shell shocked and shaggy-dog story. Inuit children may therefore have different intuitions about the notion 'word' than English children. Or again, if you ask a speaker of German how many syllables there are in a particular German word, you will typically get a quick and correct answer, but if you pose the same question to a speaker of Japanese about a Japanese word, you may well draw a blank, or get the wrong answer. This is because
syllables are not particularly salient; or, rather, another structural element, smaller than the syllable, is more salient in Japanese. This is the mora (Hyman 1985), to be discussed in chapter 9. So if you want a quick answer, you should ask how many moras there are in some Japanese word. A short vowel and each ‘half’ of a long vowel are examples of moras, such that hi ‘day’ is one mora and boo ‘stick’ is two. In addition, a consonant in the coda of the syllable is a mora. In Japanese, this could be the first ‘half’ of a long consonant (a ‘geminate’), as occurring in nattoo ‘fermented soybeans’, or a nasal consonant, like [n] as in kéndo ‘swordsmanship’. So in kéndo there are four moras: [e], [n], [o] and another [o], just as in nattoo, where [a], the first [t], [o] and another [o] are the moras.

In general, there tend to be various generalizations in a language that make reference to salient structural elements, which may in part explain why these elements are salient and thus open to the intuition of speakers (Kubozono 1999). In Japanese, for instance, there are many ways in which the moraic structure of words is relevant to the way you pronounce them and the way they are treated. However, the syllable, too, is a structural element in Japanese. For instance, it determines the possible locations of the word accent. Japanese words are idiosyncratically either accented or unaccented, and if a word is accented, an accent occurs on one of its syllables. In other words, the number of syllables determines the potential locations of the Japanese accent. Although both kéndo and bokokugo each have four moras, there are in principle four locations in which the word accent could have occurred in bokokugo (a word that happens to be unaccented), but only two in kéndo (a word that happens to have an accent on the first syllable). Given the relatively low salience of Japanese syllables, it will take a speaker of Japanese more effort to be able to say how many syllables there are in kéndo than it would take a speaker of German to do the same for the comparable German word Handy [hεndi] ‘mobile phone; cell phone’. Conversely, even though the phonology of German contains an element ‘mora’, the German speaker would probably give you a blank stare if you asked him or her how many moras there are in that word.

Q1 The lines of the lyric Do-re-mi from the 1959 musical The Sound of Music (Richard Rodgers and Oscar Hammerstein) have the same number of syllables as the notes in the song. They are given in (1). In (2), a popular Japanese version of this song is given, again with the numbers of syllables in each line.

Why do only the first and last lines of the Japanese version have seven syllables?

<table>
<thead>
<tr>
<th>(1)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Doe, a deer, a female deer</td>
<td>7</td>
</tr>
<tr>
<td>Ray, a drop of golden sun</td>
<td>7</td>
</tr>
<tr>
<td>Mi, a name I call myself</td>
<td>7</td>
</tr>
<tr>
<td>Far, a long, long way to run</td>
<td>7</td>
</tr>
<tr>
<td>Sew, a needle pulling thread</td>
<td>7</td>
</tr>
<tr>
<td>La, a note to follow sew</td>
<td>7</td>
</tr>
<tr>
<td>Tea, I drink with jam and bread</td>
<td>7</td>
</tr>
</tbody>
</table>
1.2.3 The role of spelling

Above, we saw that awareness of segmental structure can be induced by cultural factors like learning an alphabetic writing system. At the same time, however, a language’s orthography can be a hindrance to understanding the structure of language, because of the irregularities in the relations between letters and segments. The English word laugh has five letters but three segments, [l], [əː] or [æ] depending on the dialect, and [f], while the French word taxi has four letters and five segments, [t], [a], [k], [s] and [i]. Speakers of such languages will tend to mix up letters and segments in their attempts to become aware of the pronunciation of words. This is the reason why segments have been assigned unique symbols by the International Phonetic Association (IPA), which are generally used by linguists. They are written in square brackets, as has been done here. The IPA has a website featuring all these symbols, many of which were taken from the Roman alphabet, and their meanings, with sound files giving examples.²

<table>
<thead>
<tr>
<th>(2)</th>
<th>Do wa donatu no do</th>
<th>7</th>
<th>‘Do is do of donatu (donuts)’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Re wa remon no re</td>
<td>6</td>
<td>‘Re is re of remon (lemon)’</td>
<td></td>
</tr>
<tr>
<td>Mi wa minna no mi</td>
<td>6</td>
<td>‘Mi is mi of minna (everyone)’</td>
<td></td>
</tr>
<tr>
<td>Fa wa faito no fa</td>
<td>6</td>
<td>‘Fa is fa of faito (fight)’</td>
<td></td>
</tr>
<tr>
<td>So wa aoi sora</td>
<td>6</td>
<td>‘So is so of the blue sora (sky)’</td>
<td></td>
</tr>
<tr>
<td>Ra wa rappa no ra</td>
<td>6</td>
<td>‘Ra is ra of rappa (trumpet)’</td>
<td></td>
</tr>
<tr>
<td>Si wa siawase yo</td>
<td>7</td>
<td>‘Si is si of siawase (happy)’</td>
<td></td>
</tr>
</tbody>
</table>

Q2 In one of his sketches, the American comedian and singer Tom Lehrer referred to an eccentric friend of his called [ˈhenri], who spelled his name ‘h’, ‘e’, ‘n’, ‘3’, ‘r’, ‘y’. After allowing his audience to ponder the merits of the spelling Hen3ry for a couple of seconds, he added: ‘The “3” is silent, you know’.

1 What is meant by a ‘silent letter’?
2 Argue on the basis of the English word light that ‘silent letter’ is not a particularly helpful notion.

1.3 WHAT LINGUISTS DO

As a group, linguists try to become aware of the entire structure of the languages they study, even if each individual linguist is usually concerned with a specific aspect of those structures. The difference between the incipient awareness that lay language users have – which apart from being hopelessly incomplete is often grossly
mistaken – and the work of linguists is that the latter’s efforts take place in a scientific context. Thus, they will typically give evidence why they assume some structure to exist and will hold themselves accountable for mistakes by publishing their hypotheses in peer-reviewed journals.

In principle, linguists aim to discover the entire structural system of every language, motivating hypothesized structures as they go along. Of course, this goal is unrealistic, because languages are dying out at a rapid rate and the number of linguists falls far short of what would be needed for that task. It is hard to say how many languages would be sufficient to be able to answer the question that defines a further goal of linguistics: what factors allow linguistic structures to exist, or, rather, what factors account for the distribution of those structures across the languages of the world? Why are some elements common and others rare, and yet other imaginable ones apparently non-existent? Those factors will come from many different domains, like our brains (cognitive factors), our perceptual system and our speech production system (together, these are phonetic factors), as well as social factors. To go further still, how do these factors explain the way that languages are acquired by infants? And what is the explanation for the finding that from one generation to the next the linguistic structures that are acquired are slightly different, such that over time languages change in a way that makes it hard to see that stages of a language that are a thousand years apart are actually ‘the same language’?

1.3.1 Language-internal evidence

The evidence linguists present is of two kinds: language-internal and language-external. In the case of language-internal evidence, the assumptions that linguists make about some element in the language structure are backed up by regularities in the language that refer to the hypothesized element. We have seen an example of this above: the syllable determines where the accent can come in Japanese. To repeat this point, a disyllabic structure like [hasi] could be hási, hasí or hasì (recall that a Japanese word may be unaccented), and these actually happen to be words: ‘chopsticks’, ‘bridge’ and ‘edge’, respectively. By contrast, there are only two theoretically possible words that consist of the syllable [ree]: rée and ree, which again are real words, ‘a bow in greeting’ and ‘gratitude’, respectively. It doesn’t make sense to ask what reé might mean in Japanese: it couldn’t be a word, because the accent never occurs on half of a syllable. Now, for it to be the case that the number of accentable positions is equal to the number of syllables it must of course be the case that the syllable exists.

Q3

1 How many Japanese words could be formed from [hi]? And how many words could be formed from [kentoo]?
2 There are, in fact, two Japanese words that are pronounced [rée]: ‘example’ and ‘bow, greeting’. What do you call words that have the same pronunciation?
1.3.2 Language-external evidence

In addition to the language-internal evidence illustrated with the example of the Japanese syllable, there are many other sources of evidence. Most importantly, linguists can make phonetic measurements. They can measure the acoustic properties of speech, track the articulatory gestures by speakers and register listeners’ responses to speech stimuli in perception experiments. These measurements enable them to understand better just what is going on when people communicate through spoken language. Such phonetic research not only reveals in detail the physiological, physical and psychological events that together make up the ‘speech chain’, but can also be used to provide evidence for or against a particular structural hypothesis. For instance, to show that the mora is an important element in the structure of Japanese, we could collect Japanese words varying in the number of syllables and the number of moras. One category would have words of one mora, like go ‘word, language’; another words with two moras and one syllable, like kúu ‘bite, eat (by insects)’; a third words with two moras and two syllables, like motí; and so on. If we then measured the durations of the words, we could see whether word duration is better explained by how many moras a word has or by how many syllables a word has. If it turned out that the mora explained better than the syllable how long a word is, we would have phonetic evidence for the greater relevance of the mora in Japanese.

Behavioural experiments try to find evidence for linguistic structure by registering subjects’ responses to specifically designed tasks. Such tasks should, of course, never directly ask subjects to give answers to questions about linguistic structures, or at least never take such data at face value, because this carries the obvious risk of including all sorts of misconceptions in the evidence. To return to language art forms, we could use the rhythm of poetry to investigate the syllabic structure of words. If we wanted to know how many syllables there are in an English word like hire, we could ask whether names like Hire Fire and Mo Toe could take the place of Humpty-Dumpty in the nursery rhyme Humpty-Dumpty sat on a wall. If all is well, there would be significantly more votes for Hire Fire than for Mo Toe, which difference could then be argued to be due to the fact that hire and fire are disyllabic words, like Humpty and Dumpty, while Mo and Toe are not. As you may have guessed, real-life experimental questions and tasks are more complex than these examples suggest.

The research agenda of psychologists, whose interest in language received a boost in the 1960s, was to unravel the mental processes that allow listeners to understand speech and the processes that allow speakers to produce it. You will be familiar with the sensation you have that when listening to a foreign language you cannot tell where the words are. The reason for this is of course that, within a coherent phrase, there are no pauses between the words, any more than there are pauses between the syllables or moras of words (Cutler 2011). In fact, if by ‘pauses’ we were to think of brief periods of silence, then we will find quite a number of ‘pauses’ in coherent phrases, but these will occur within words as well as at the edges of words with reckless abandon. In an English phrase like paper napkins, there are two such silent
intervals, one corresponding to the second [p] in paper and one to the [pk] combination in napkins. The boundary between the two words does not correspond to any silent interval at all. If there are no audible word beginnings, the question arises how native speakers of English avoid hearing words made up of sound sequences inside words, like pay in paper or nap in napkin, or even across word boundaries. Figure 1.1 shows a speech waveform of this phrase, with lines indicating the beginning and ending of sounds, as well as the word boundary. In this case, we cannot even say where the word begins. This could be defined as the moment that the lips close for the formation of [p], but since this action produces no acoustic energy, it cannot be distinguished from the silence that occurred before this event.

In fact, psycholinguists have shown that during speech perception, listeners do hear strings of sounds inside words as possible words, but this happens very briefly and below our awareness. Word strings across word boundaries may also ‘activate’ words that were never said, such as income in mustering compassion (. . . muster income passion . . . ), but because of the particular ways in which sounds combine into words, cross-word strings of sounds are less likely to constitute words. In fact, languages vary in the way their speakers detect word boundaries, depending on the phonological regularities that apply to word boundaries (e.g. all words may begin with a stressed syllable, as in Finnish) and the phonetics (e.g. all vowel-initial words are provided with a glottal stop before them, as in German).

Q4 Why would the English phrase Call the police! never cause a word to be heard made up of a cross-word string of segments, while this may well happen in Wildest endings?

Other research has shown that it is easier to think of a word, as measured by the time it takes to say the word for an object in a picture, when the speaker has just
been alerted to some other word that begins in the same way. That is, after just hearing checkers, people are quicker to say 'Cherry!' when presented with a picture of a cherry than after just hearing backgammon. Priming research of this sort can go some way towards understanding how speakers retrieve the pronunciation of a word before they actually pronounce it (Meyer and Belke 2007).

Psycholinguistic research into the processing of language in the brains of speakers and hearers and, more recently, neurological research into brain activity during language processing should of course be informed by the results of linguistic research into the structures of languages. Equally, structural linguistic research should take place in the wider context of psychological, neurological and sociological research, something that has not always been the case. In fact, some people would say that linguists are notorious for working out their problems in isolation from other fields, while some linguists may feel that psycholinguistic research has not yet addressed the possible roles of all aspects of the linguistic representation.

1.4 MORPHOSYNTACTIC STRUCTURE

Phonology is concerned with a particular aspect of linguistic structure. In order to see what part of linguistic structure is phonological, we will first briefly consider that part of the linguistic structure which is not phonological, the morphosyntactic structure. The morphosyntactic structure of a language can be seen in the arrangement of the meaningful units of any linguistic expression. A distinction is made between morphology, which deals with the structure up to the level of the word, and syntax, which deals with the structure above the word.

1.4.1 Morphemes and words

The smallest morphosyntactic unit is known as the morpheme. The word scratch is a single morpheme, but the word pens contains two morphemes. A morpheme that can be a word by itself is a free morpheme. The morpheme pen could form a word by itself, and is for that reason an example of a free morpheme, while the morpheme s is a bound morpheme, because it must combine with some other morpheme in order to be part of a word. A word consisting of a single (necessarily free) morpheme is a simplex word, other words being complex. An important group of bound morphemes are affixes, which can be divided into prefixes and suffixes, depending on whether they are placed before or after what is known as the base. Thus, the s in our example is a suffix because it is placed after the base pen, and so is able in the complex word scratchable.

Most words have a category membership, which allows them to be used in certain positions in sentence structure. If we assume that a sentence consists of a Noun and a Verb, in that order, then the sentence Pens scratch is well-formed, while *Scratch pens is an ill-formed sentence (as indicated by the *) if pens is a subject. Affixes are subcategorized for a category membership, meaning that they only attach to bases that are members of a particular category. Thus, the plural suffix s
attaches to nouns, while the suffix *able attaches to transitive verbs, i.e. verbs that take a direct object, such as *scratch, *read and *drink (but not *sleep: *sleepable). Affixes are also distinguished on the basis of whether they create new words, in which case they are known as derivational affixes, or merely create new forms of the same word, like the plural suffix *s or the past tense suffix in *scratched, in which case they are inflectional. Derivational suffixes come with their own category membership, which is frequently different from the category of their base. Thus, *able has the category Adjective, and every complex word ending in *able is therefore an Adjective. In (1), the structural information for these two affixes is given in a notation known as labelled bracketing.

(1) a $[[ N ]_z] \text{ PLURAL}$
   b $[[ V_{trans} \bar{abl}]_{adj}}$ ‘is able to *Verb’

Observe that a base can be simplex or complex. In *scratchers, the plural suffix was attached to *scratcher, itself composed of *scratch and er. A simplex base which is a free morpheme is also known as a stem, while a simplex base that is not, like the English verbal base mit in transmit, is also known as a root.

Q5 Explain why there are no words *dieable and *seemable in English, and why imaginable is a word.

1.4.2 Syntactic structure

Words do not directly enter into sentence structure. If the structure of our simple sentence $S$ were simply $[N \ V]$, it would only be possible to form sentences of the type Kittens *scratch, Pens leak and Time *flies. In fact, we can also form A good pen mustn’t *scratch or The pen I bought in Italy *leaks. Here, the structural position of *Pens in our original example sentence *Pens *scratch is occupied by the word groups A good pen and The pen I bought in Italy. This is the reason why, in between the structural level of the word, illustrated in (1), and the structural level of the sentence (or clause), we need to recognize the structural level of the (syntactic) phrase. That is, instead of $[N \ V]$, our simple sentence has the structure $[Noun Phrase – Verb Phrase]$, or $[NP \ VP]$. In addition to the NP and the VP, another frequent type of phrase is the Prepositional Phrase, or PP, which consists of a preposition plus an NP, such as with the leaky pen, on the surface or in Italy.

The introduction of a level of phrasal structure allows us to point to two properties of morphosyntactic structure. One is that this structure is hierarchical. This means that we can distinguish higher and lower levels of structure: a word level, a phrase level and a sentence level. The second property is recursiveness. This means that a unit of a given structural level can be incorporated into a unit of the same structural level, or that a unit of a higher level of structure can be incorporated...
into a unit of a lower level. For example, the sentence *The pen I bought in Italy leaks* consists of two phrases, the NP *The pen* and an S (*which I bought in Italy*), which sentence in turn consists of the NP *I* and a VP *bought which (i.e. the pen) in Italy*. The structure in (2) expresses this. In (3), we give the mini-grammar that will produce this sentence. Recursiveness is shown in (3b), first, by the fact that we can feed a higher-level S into a lower-level NP and, second, by the fact that we can feed an NP into an NP. As you will have guessed, we can in principle apply the same rules again and again. It may be fun to do this, as illustrated by the children’s verse *This is the farmer that kissed the girl that chased the cat that killed the mouse*, etc.

\[(2) \quad |S|_{NP} \{NP \{NP \{The \pen\} \PP \{bought \,(sc. \,the \pen) \,in \,Italy\}\}|_{VP}\{leaks\}|_{VP}\{S\}\]

\[(3) \quad a \quad |S|_{NP} \quad \quad \quad \quad |NP|_{S} \quad \quad \quad \quad |VP|_{NP}\{bought\}|_{VP}\{leaks\}|_{VP}\{S\}\]

There is much that is the same in the syntax of the languages of the world, but there is also a good deal of variation. Some languages have a VP in which the Verb precedes the rest (i.e. [V NP] or [V PP]), while other languages maintain the opposite order (i.e. [NP V] and [PP V]). Moreover, some languages have a syntactic rule that moves the Verb to initial position in the sentence. This type of variation is often referred to as that between S(ubject)V(erb)O(bject), SOV and VSO.

Q6 Identify the nouns and verbs in the following words:

1. *unscratchable*
2. *road tax increase*

Q7 Identify all the NPs in the following sentence:

*He wrote the letter with a pen he bought in Italy*

1.5 A WORLD WITHOUT PHONOLOGICAL STRUCTURE

Spoken languages use human vocal sound to give shape to their morphemes. Phonology is the branch of linguistics that aims to describe the way in which this medium of human vocal sound is structured, in languages generally as well as in individual languages. To see what is meant by the sound structure of languages, it may be instructive to pretend for a while that languages do not have it. Imagine that every morpheme of a language were assigned some vocalization. Conceivably, these vocalizations could be quite lengthy in view of the large number that would
be needed to distinguish all the morphemes of the language. But, importantly, there would be no implication that they should have structure in the sense that they are composed of subparts, any more than abstract paintings are. To give a hypothetical example, vocalization (4) might be the morpheme meaning 'oak tree' in some language.

(4) 'oak tree': high-pitched wheeze, trailing off into a voiced cough with central, nasal vowel quality

The wide variation in morphosyntactic structure that is found in the languages of the world could exist without there being any structure to the human vocal sounds that languages use as a medium. Clearly, in the hypothetical situation described above, languages could still have SVO or VSO as their basic word order. They could either have an extensive Case system (i.e. be strongly inflectional) or have many prepositions; they may or may not have articles, they may or may not mark plural in both Subjects and Verbs ('Concord') and so on. Where our hypothetical situation is different from the situation in the real world is that nothing would have to be said about the sound structure, because there would not be any. All that anyone could do is make a list of descriptions like the one in (4).

Q8 Why, in the hypothetical situation above, would it be impossible for speech errors like [klɪs keə] for kiss Claire to occur?

Now let us turn to real life. The first observation to be made about the pronunciation of morphemes is that there are recurring elements. For instance, the sound patterns used for morphemes can be analysed as strings of segments. This is the basis on which we can say that English cat has the same segments as tack or act, although the segments occur in different orders. Minimally, then, the task of phonology is to state what these recurring elements are.

The second observation to be made is that the recurring elements do not occur in all possible orders. For example, while cat, act, tack are possible combinations in English of the segments [k], [æ] and [t], the same does not go for [æt[k]. Notice that it is not the case that the sequence [æt[k] never occurs in English. This sequence is part of the word Atkins, for instance, and occurs in the sentence The cat killed the mouse. What this shows is that there is some constituent higher than the segment which imposes constraints on what sequences of segments it may contain. In the example, that constituent is the syllable. That is, *[æt[k] cannot occur as a sequence of segments inside the syllable. In general, elements at one level of structure combine in restricted ways to form elements of a higher level of structure. We will see in chapter 5 that the segment is to be looked upon as an element that itself combines elements of a lower structural level, called distinctive features. So a second task for phonological theory is to state what the permitted patterns of arrangement of the phonological elements are.
A third important observation is that segments may be pronounced differently depending on their environment. This is because languages usually have processes that affect (‘change’) segments in particular contexts. In English, for example, [p t k] are aspirated when they are the first consonant of a syllable. Hence, tack is pronounced [tæk], but stack is [stæk]. Similarly, the Dutch [z] in [ze:] ‘sea’ is pronounced [s] when it is preceded by an obstruent, as in [ɔp se:] ‘at sea’, but not elsewhere, as in [a:n ze:] ‘at the seaside’. Or again, in French the final consonant of an adjective like [pətit] ‘small’ is pronounced in le petit autobus ‘the little bus’, but not in le petit camion ‘the little lorry’, as a result of the difference in the first segment in the following noun. As these examples make clear, one and the same morpheme may have different pronunciations in different contexts as a result of the existence of these processes.

Minimally, then, a phonological description will have to answer three questions: what are the phonological elements, how do they combine into higher structural elements, and what generalizations exist about their contextual variation?

Q9 Do you think that sign languages for the deaf have phonologies?

1.5.1 One phonology for all languages?

Now, it might have been the case that all human languages had the same sound structure. In order to see what this would mean, let’s once more enter into an imaginary world. We will make up a ‘universal phonology’ and then give partial descriptions of two – very different – languages. The universal sound structure is given in (5), and (6) gives partial descriptions of the two languages.

(5) a Segment inventory:

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>p t k</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>b d g</td>
<td>a</td>
<td></td>
</tr>
<tr>
<td>m n l</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b Arrangement:
A word consists of one, two or three syllables.
A syllable consists of CV.

c Process:
FRICATION Initial obstruents in non-initial words are fricatives.

Because of frication, [p t k] will be [f θ x], and [b d g] will be [v ð ɣ], in the context specified. Because the fricatives are variants of the plosives, they are not listed separately in (5a). Observe that the labial obstruent in this language has two pronunciations, [p] and [f]. The distinction between a sound category or phoneme, here /p/, and its different pronunciations is usually expressed by writing
the phoneme between slashes and the pronunciations between brackets, [p] and [f] (see chapter 7).

Using the data in the first column of (6), we find that in Language I the sentence meaning ‘Cows are grazing in the field’ would be [pilu namu ma ðiku]. Here, Subject-Verb Concord is expressed by the double occurrence of [lu], while the position of the PP ‘in the field’ is that of the O in SVO.

Q10

1  Why does the word for ‘field’ in the sentence of Language I begin with [ð], and not with [d]?
2  What is the translation of this sentence into Language II?
3  Why would the process we assumed in (5c) be a convenient feature of human language, as seen from the point of view of the listener?

The reason why a single phonology for all languages is so improbable is that the phonologies of languages change over time, just as does the morphosyntax. One general factor inducing change affects both aspects of structure. When learning their language, infants may make different generalizations based on the data from what their parents did when they learnt the language. Changes induced by such generalizations will affect word order as well as the forms of specific words. To begin with a syntactic example, the position of the verb changed from an earlier SOV in Proto-Germanic to SVO in English. Other Germanic languages remained SOV, except that they moved the finite verb form to the second position in main clauses. As a result, English now has SVO in both dependent clauses (7a) and main clauses (7b), but Dutch has SOV in dependent clauses and a partial SVO, known as ‘Verb-Second’, in main clauses: $S-V_{\text{finite}}-O-V_{\text{nonfinite}}$ (Koster 1975). Here, English acquired a more general rule than did Dutch and other Germanic languages.

(7)  

<table>
<thead>
<tr>
<th>English</th>
<th>Dutch</th>
</tr>
</thead>
<tbody>
<tr>
<td>a  I think she would like to buy fish</td>
<td>Ik denk dat ze vis wil kopen</td>
</tr>
<tr>
<td>b  She would like to buy fish</td>
<td>Ze wil vis kopen</td>
</tr>
</tbody>
</table>
A morphological example concerns the Dutch 3sg verbal suffix [t]. It appears after stems like *ken* ‘know’ and *raak* ‘hit’, which thus change to *kent* and *raakt* when used with a 3sg subject NP. However, auxiliary verbs like *wil* ‘would like’ and *kan* ‘can’ don’t take this [t]: *De jongen kent het* ‘The boy knows it’, but *De jongen kan het* ‘The boy can (do) it’. Children often generalize t-suffixation to the auxiliary verb *wil* and say *De jongen wilt het* ‘The boy wants it’. In this case, the treatment of *wil* is probably inspired by the meaning ‘want to’, which is expressed by a non-auxiliary verb in many languages, even though Dutch *wil* behaves just like the other auxiliaries in other respects. Such novel generalizations are usually modified in later stages of the acquisition process, as when children acquiring English start using *mice* instead of the *mouses* which they may have produced before, but some stay and lead to language change.

A second factor relies on changes in the pronunciation of specific sounds, or of specific sounds in specific contexts. Such changes are very common, although we may not always realize that the variation we observe between speakers may be an indication of change. Such changes are sociologically determined in the sense that they arise within a smaller group of speakers. Again, some of these changes will become more general, as happened in the case of the disappearance of [r] from the ends of syllables in a number of varieties of English. In the English spoken in England, this disappearance caused *court* to have the same pronunciation as *caught*, and *garter* to rhyme with *sonata*. As a result, the phonology of the innovating ‘r-less’ speakers represents a fairly drastic change relative to that of the older system, for instance in having additional, new vowels, like the diphthong [ɪə] in a word like *beer*. In the English spoken in England, *beer* now rhymes with *idea*, but in the English spoken in the USA and Canada, which preserves an ‘r-full’ version of the language, they do not rhyme.

### 1.6 Two kinds of structure

The fact that the vocalizations that are used to represent the meaningful units of language themselves have structure has an important consequence, which is that terms like ‘constituent’, ‘element’ and ‘unit’ are ambiguous. The speaker or writer might be referring to a morphosyntactic unit, such as ‘the morpheme *hill*’ or ‘the sentence *I like it*’, or to a phonological unit like ‘the segment [ɛ]’ or ‘the syllable [kɛt]’.

<table>
<thead>
<tr>
<th>Q11 Divide the following English words up into (a) morphemes and (b) syllables:</th>
</tr>
</thead>
<tbody>
<tr>
<td>elephants</td>
</tr>
<tr>
<td>palm oil</td>
</tr>
<tr>
<td>unsettling</td>
</tr>
</tbody>
</table>
Thus, a linguistic expression always has two structures, a morphosyntactic one, which reflects the meaningful elements in the expression, and a phonological one, which is the structure most immediately relevant to the pronunciation of the expression. The distinction is a very real one, because the morphosyntactic constituents do not map one to one onto the phonological constituents; to use a technical term, they are not isomorphic. That is, morphemes do not exclusively correspond to segments, or exclusively to syllables, etc. While the Dutch polite second person pronoun consists of the one segment [y], the morpheme [ɪk] ‘I’ consists of two. And while these two pronouns each consist of one syllable, the informal second person plural pronoun consists of two: [jɪy.li]. (The period ‘.’ is used to indicate syllable boundaries.) In English, the single syllable [suːz] can represent two morphemes, Sue and is, as pronounced in the sentence Sue is ticklish. In (8), the morphosyntactic structure of that sentence is given. The way the sentence is analysed in morphosyntactic constituents should be compared with the (partial) phonological constituent analysis of that same sentence, given in (9). Notice that just like the morphosyntactic structure, the phonological structure is hierarchical. There is a layer of segments, which build a layer of syllables (symbolized σ), which build a layer of feet (symbolized F), which build a layer of phonological words (symbolized ω), and so on. In addition to the different constituent structures for Sue is, for example, notice that ticklish consists of two syllables as well as of two morphemes, but that the syllables and the morphemes do not divide the word up in the same way. The dual structure outlined in this section exists in all languages. It is dealt with in more detail in chapter 12.

Q12

1. In the case of English [æ], [k] and [t], all permitted arrangements are in fact words: cat, tack, act. But if we replace [æ] with [i], combining the three segments yields only two words: kit and tick. Would it be correct to say that [ikt] is an ill-formed combination in English?

2. Can you explain why [mrɛk] would never be introduced as a brand name in English, while [krɛm] might well be?
1.7 CONCLUSION

This chapter has made the point that linguistic expressions have two parallel hierarchical structures. One of these, the morphosyntactic structure, reflects the meaning of the linguistic expression, while the other, the phonological structure, reflects its pronunciation. A given constituent in either structure will typically not consistently map onto any single constituent in the other structure, a point that will be worked out more in chapter 12. Before we move on to a further discussion of the phonological structure, we first deal with the way we produce and articulate speech, in chapter 2.

NOTES

2 http://web.uvic.ca/ling/resources/ipa/charts/IPAlab/IPAlab.htm
2.1 INTRODUCTION

To describe how we produce speech and what speech looks like acoustically in the space of a single chapter is a tall order, and you would do well to consult other textbooks that deal more specifically with the phonetics of speech, like Catford (1988), Laver (1994), Ladefoged and Maddieson (1996), Ladefoged and Johnston (2010), Ashby (2011), Reetz and Jongman (2011) or Zsiga (2012). We describe the speech production process in two stages. First, we consider the role of the lungs and the larynx. This part of the speech organs is responsible for the actions of the vocal folds, which are located inside the larynx. A common and spectacular action of the vocal folds occurs when they vibrate against each other so as to produce a buzzing sound, which can be varied in pitch and loudness. Second, we deal with the role of the channel extending from the larynx onwards, called the vocal tract. It is formed by the pharynx, the mouth and, for nasal and nasalized sounds, the nasal cavity. The vocal tract modifies the buzzing larynx sound. In its unmodified form, it would sound much like a small petrol engine, but this is hard to verify so long as we keep our heads on! Because the vocal tract can assume many different shapes, these modifications are highly varied. The most striking effect here is the production of different vowel sounds. The term organs of speech is used to refer to parts of the body in the larynx and the vocal tract that are involved in the production of speech. It is a misleading term in that it suggests that we have special physical organs for speaking. This is not so: all our so-called organs of speech have primary biological functions relating to our respiratory system and the processing of food.

The pronunciation of words is conventionally represented with the help of phonetic symbols, any such representation being a phonetic transcription. The symbols used in this book are those proposed by the International Phonetic Association (IPA), which can be found on page ix. A phonetic symbol stands for a particular speech sound, or segment, which is defined independently of any language. Phonetic symbols may be accompanied by diacritics, signs which are printed above or below a phonetic symbol or with which the symbol is superscripted, and which specify particular features of pronunciation. For example, in the transcription [kʰæt], which indicates the pronunciation of the English word cat, [kʰ] represents a [k] which is accompanied by aspiration, a brief [h]-like sound occurring between the [k] proper and the following [æ]. It is not always necessary, or even desirable, to indicate all the features of the pronunciation of a word in a transcription: the transcription [kæt] is often sufficiently informative if the reader knows the language concerned. A transcription that includes a great deal of detail is called narrow.
2.2 THE LUNGS AND THE LARYNX

A crucial requirement for the production of acoustic energy is a mechanism to create an air pressure difference in the appropriate locations in the larynx and the vocal tract. There are a number of ways in which air pressure differences for speech production are created, called airstream mechanisms (Abercrombie 1967). The most commonly used by far is the pulmonic airstream mechanism, and the great majority of speech sounds are produced with the help of increased air pressure created by our lungs. Before we begin to speak, we breathe in, taking in sufficient air to produce an utterance of reasonable length. Instead of simply letting go of the muscular tension and allowing our lungs to collapse, pushing the air from them (which is what we would do if we were breathing normally), we slowly ease up on the tension, thereby slowing down the exhalation phase. This artificially extended period of pressure from our lungs is used to produce speech. Because it is the exhalation phase rather than the inhalation phase that is used, these speech sounds are called egressive. Most languages only have pulmonic egressive sounds. In section 2.8, we will briefly describe the production mechanisms of three types of nonpulmonic sounds (clicks, implosives and ejectives).

After passing through the bronchi and the trachea, the first organ the airstream will meet on its path from the lungs is the larynx. The outward part of this organ can be felt – and, especially in men, be seen – at the front of the neck (the Adam’s apple). The larynx is a valve, which can be opened and closed by two thickish flaps that run from back to front inside the larynx (see Figure 2.1). These flaps are primarily there to prevent food or saliva from entering the lungs, but because they also have a function in speech they are known as the vocal folds or the vocal cords. The aperture between them is called the glottis. No air can pass through the glottis when it is closed, while the air can flow quite freely through an open glottis.

Figure 2.1
Schematic drawings of (a) a closed glottis, as during the closure stage of a glottal stop or during the closed phase of the vibrating glottis; (b) an open glottis; (c) a narrowed glottis.
The pulmonically produced pressure difference is used for three purposes. First, it can be used to drive the vibratory opening and closing actions of organs like the vocal folds, as explained in section 2.2.1. During the articulation of [m], for instance, vocal-fold vibration can easily be felt by placing one’s fingertips on one’s ‘Adam’s apple’. Second, it can be used to generate a flow of air that can be channelled through a narrow opening to create audible air turbulence, or friction. This happens during the articulation of [s], for instance. Third, it can be used to build up pressure behind a complete blockage of the vocal tract in order to create an explosive sound when the blockage is suddenly removed. This occurs in the articulation of plosive consonants like [p], such as when we say [pa].

2.2.1 The vocal folds: the open and vibrating glottis

There are many consonants that are produced with the glottis held open, as in ordinary breathing. Such sounds are called voiceless, and we hear them because other speech organs, the tongue or the lips, are used to generate fricative or explosive sounds further up in the vocal tract. Examples of voiceless sounds are [f] and [ʃ] in fish and the sound sequence [st] in stay.

The vocal folds are exploited in various ways to create sound which can be used as a basis for speech. This is known as phonation. The most important type of phonation is voice, which is produced when the vocal folds vibrate. Vocal-fold vibration occurs when the closed glottis is subjected to increased subglottal air pressure which is sufficient to blow the vocal folds apart, but not enough to prevent them from falling together again when the air pressure between them drops as a result of the Bernoulli effect, a physical effect which causes pressure minima at points where the flow of gases or liquids is high. As soon as they have been sucked together, the vocal folds are once more blown apart as a result of the subglottal air pressure. This process typically repeats itself more than 100 times per second for the larger and laxer vocal folds of men, and over 200 times per second for the smaller vocal folds of women. Consonants like [m], [l] and [j] (which we will see are sonorant consonants) and vowels are normally voiced. Voiced obstruent consonants also exist, such as [b] in English abbey and [z] in lazy, and usually contrast with their voiceless counterparts (in these cases with [p] and [s], as in happy, lacy).
2.2.2 Devoicing and aspiration

When a consonant that is normally voiced is pronounced without vocal-fold vibration in some context, it is said to be **devoiced**. Devoiced segments are symbolized with a circle below the symbol (which may also appear above it, if there is no space below it). For example, devoicing may follow voiceless obstruents, in particular plosives. In many languages, the vocal folds may begin to vibrate immediately after the release of the closure made for the plosive for a following voiced segment, as in [pa], but in other languages the vocal folds may remain open for a while. In the latter case, the plosives are said to be **aspirated**. As shown above, aspirated plosives are symbolized with a superscript [h], as in [pʰa]. English has voiceless aspirated plosives at the beginning of the syllable, as in tea, pea, key. If a sonorant consonant rather than a vowel follows the aspirated plosive, the aspiration is indicated by means of the devoicing diacritic, as in English [pʰei] play.

The timing relation between plosive releases and the onset of vocal-fold vibration is expressed as **voice onset time**, or VOT. The VOT is zero when the plosive release and the onset of vocal-fold vibration are simultaneous; when the onset of vocal-fold vibration is earlier than the release of the plosive, VOT is negative, and when it is later, it is positive. Typical values for aspirated plosives are between VOT +50 ms and VOT +80 ms.

Figure 2.2 shows a **speech waveform** of the word pass, spoken by a speaker of British English as [pɑːs]. The first acoustic event is the burst of [p]. (Recall from chapter 1 that we cannot tell from the acoustics when a [p] begins if it occurs after silence.) Then there is some weak turbulence, which is the aspiration as indicated by [h]. The VOT is measured from the beginning of the burst to the beginning of the vowel, and is 61 ms in this example. The waveform during the vowel shows the vibrating vocal folds. Finally, the voiceless [s] consists of loud turbulence. Voiced sounds like [m] and [ɑː] are ‘periodic sounds,’ because the waveform shows a

---

**Figure 2.2**
Speech waveform of the British English word pass. The vertical lines demarcate the release of the bilabial [p], the beginning of the vocal-fold vibration for [ɑː], the beginning of [s] and the end of [s]. The voice onset time (VOT) is given by the duration between the first two demarcation lines.
The production of speech

repeating pattern. Each of the repeated portions of the waveform corresponds to an opening-and-closing action of the vocal folds and is known as a period. Figure 2.3 reproduces a section from the waveform in Figure 2.2, with one period marked out. Notice that the shape of the waveform during this period is more or less the same as those of the periods before and after. The duration of the period, usually measured in milliseconds (ms), depends on the frequency with which the vocal folds carry out their opening-and-closing actions. If these occur 200 times per 1,000 ms (expressed as 200 Hz), the period will be exactly 5 ms. The frequency of vibration of the vocal folds is the fundamental frequency of the speech signal, also referred to as the f0 (‘F-zero’). The shape of the waveform during a period determines the quality of the sound, in this case that of the British English vowel [æː]. Panel (b) shows a period of about the same duration for the vowel [iː] in peace, and panel (c) does the same for [ʊ] in foot. The different shapes of the waveforms are due to the different shapes the vocal tract assumes during the production of these vowels: the resonances that are set up in the air in the vocal tract depend on its length and shape. The relation between these resonance frequencies, known as formants, and vowel qualities is briefly discussed in section 2.4.

The turbulent signal for [s] has no fundamental frequency. It consists of ‘noise’, a largely random pattern of vibrations in which broad frequency zones are more
emphasized than others. These different emphases in the frequency spectrum determine whether the turbulence sounds more like a high-friction [s] or a low-friction [f], etc. When turbulence and vocal-fold vibration are produced simultaneously, we produce sounds like [z] and [v].

**Q14** What would the speech waveform of the word *ceased* look like?

**Q15** What is the approximate fundamental frequency of the vowel shown in panel c of Figure 2.3?

### 2.2.3 Special types of phonation

Three special types of phonation are mentioned here: whisper, breathy voice and creaky voice.

1. The vocal folds can be brought together to form a **narrowing** which produces **friction** when air passes through it. This is how people whisper: instead of voice, glottal friction may be used as the acoustic source to be modified. A whispered speech sound occurring in otherwise voice-phonated speech is [h], as in English [hæt] *hat*. Whisper can be indicated by the devoicing symbol, as has been done in [æ̥æt].

2. **Breathy voice** occurs when the closing phase of the vibration is not complete, so that air is allowed to flow through with friction during phonation. Breathy voice is used in European languages to signal confidentiality (Laver 1994: 200), and is sometimes used to create the effect of a sexy voice. Voiced aspirated plosives, which occur in many languages spoken on the Indian subcontinent, are produced by allowing breathy voice to be used throughout the plosive and the following vowel, as in Hindi [bɦal] ‘forehead’. Breathy voice can be indicated by [\̥], as in Dutch [u̥t] ‘hat’, alternatively transcribed as [fiut].

3. **Creaky voice** or **laryngealized voice** is produced with tight vocal folds, and often allows the listener to hear the opening actions of the vocal folds as separate events. (The effect may remind you of the sound produced when running a fingernail across the teeth of a comb.) British English speakers may break into creak at the ends of their utterances, when the pitch is low. Both male and female speakers of American English may have consistent creak during longer parts of utterances, together with fairly low pitch. Many Nilotic languages use laryngeal voice contrastively (Ladefoged 1971). Dinka has a set of vowels with creaky voice, symbolized by means of [\̰], which contrast with a set with breathy voice, as in [r̩o:r] ‘forest’ – [r̩o:r] ‘men’ (Andersen 1987).
2.2.4 Pitch

Variations in the frequency of vibration are heard by the listener as variations of pitch: the more frequently the vocal folds open and close, the higher the pitch. In languages like English, French, German and Spanish, variations in pitch are used to signal different discoursal meanings. For instance, in the utterance But I don't want it!, the syllable want will be higher than the syllable it, but in Want it?, it will be the higher syllable. These two intonation patterns are known as the ‘fall’ and the ‘rise’, and add a ‘declarative’ and ‘interrogative’ meaning to the utterance, respectively. In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.

In other languages, called tone languages, different pitch patterns are used in the same way vowels and consonants are used in all languages, i.e. to distinguish.
The production of speech

2.3 THE VOCAL TRACT

The vocal tract extends all the way to the lips (see Figure 2.4). It consists of the pharynx and the mouth, to which an extra tube extending to the nostrils may be coupled, the nasal cavity. This (potentially bifurcating) tube acts as a resonator, modifying the sound produced at the glottis. The sound produced by a vibrating glottis can be modified by changing the position of the tongue, jaw and lips so as to produce a range of different vowel qualities. Second, inside the vocal tract there are further opportunities for generating sound, which in their turn will be modified by the shape of the vocal tract in front of the sound source. For instance, the tongue can be brought up against the roof of the mouth to form a constriction that generates friction when air is passed through it, as during [s] and [x]. In this section, we describe the three parts of the vocal tract mentioned above, and in section 2.4

Figure 2.4
Cross-section of the vocal tract.
we deal with the positions that the tongue, jaw and lips may assume to produce different vowels. In section 2.5 we identify the various places at which languages make articulatory constrictions. In the same section, we will classify the types of constriction that are used.

2.3.1 The pharynx

The pharynx is the vertical part of the tube extending up from the larynx to the velum. The forward wall is formed by the root of the tongue, which faces the back wall of the pharynx.

2.3.2 The nasal cavity

The soft palate, or velum, is a valve which closes off the entrance to the nasal cavity when it is pressed up, but opens the cavity when it is allowed to hang down, as in ordinary breathing. (When we have a cold, the entrance to the nasal cavity may be blocked by mucus, which forces us to breathe through the mouth, and hence to produce only oral sounds.) The velum ends in a pear-shaped little blob of flesh, which can be seen during the speaker's articulation of [a: ]. It is called the uvula.

2.3.3 The mouth

The mouth is the most important part of the vocal tract because it is here that the most drastic modifications of its shape are achieved and the majority of the articulatory contacts are made. The roof of the mouth is formed by the soft palate, with the uvula at the extreme end, and the hard palate, which lies to the front of the soft palate. With a curled-back tongue, it is possible to feel the hard palate arching back to where the soft palate begins. Immediately behind the front teeth is the alveolar ridge, which is touched by the tongue during the pronunciation of dada; then there are the upper front teeth themselves and the upper lip. Below these parts there are the lower lip, the lower front teeth and the tongue, of which the lower lip and tongue are active articulators. The zone immediately behind the tip of the tongue is called the blade. You use it when imitating the sharp, hissing sound of a snake. Together, tip and blade are called the crown, which term is from Clements (1985). The part of the tongue opposite the hard palate is called the front, the part opposite the soft palate is called the back. The section comprising both front and back is known as the dorsum.

2.4 VOtEALS

For vowels, the vocal folds vibrate. The crown is held behind the lower teeth, while the dorsum is bunched, forming a constriction that allows a frictionless escape of the air. The lips may be rounded, causing the vocal tract to be lengthened. Vowels may have a more or less stationary configuration of the speech organs or be characterized by a trajectory of the tongue or lips. Moreover, they may be oral or nasalized.
2.4.1 Monophthongs

A vowel whose quality remains stable during its production, like [æ] in English *hat* or [u:] in German *[g]ut* ‘good’, is known as a **monophthong**. Different vowel qualities are produced by different tongue or lip positions. The location of the tongue bunch can be varied in a vertical dimension as well as a horizontal dimension. In the British English phonetic tradition, four steps are recognized in the vertical dimension: **close, close-mid, open-mid** and **open**. In the American English tradition, the three terms **high, mid** and **low** are used to distinguish three vowel heights. Horizontally, the bunch can go from **front** to **back**, with **central** being used as an in-between value. If we disregard the open vowel [a], most languages just have front unrounded and back rounded vowels. Italian, for instance, has the vowels shown in (1). The position of the lips is rounded for [u], [o] and [a], and unrounded for [i], [e], [ɛ] and [a]. (Note that, deviating from IPA conventions, authors normally use the typographically simpler symbol [a] to represent a central or central to back open unrounded vowel like [α]. We have followed this practice in this book.)

<table>
<thead>
<tr>
<th></th>
<th>Front unrounded</th>
<th>Back unrounded</th>
<th>Back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>a</td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>ɛ</td>
<td>o</td>
<td>ɔ</td>
</tr>
</tbody>
</table>

If the tongue positions used for [i], [ɛ] and [ɛ] are combined with rounded lips, the vowels [y], as in French *[l]yn* ‘moon’; [ø], as in *[pø] ‘(a) little’; and [æ], as in *[æ]=el* ‘alone’ are produced. A rounded [a] is [ɔ], and may occur in *dog* as pronounced in varieties of English spoken in England and on the east coast of the USA. Unrounded vowels with the bunch in the back or centre exist also. The unrounded counterpart of [ø] is [ʌ], that of [o] is [ɔ], and that of [u] is [u].

Rounded front vowels are somewhat more central than unrounded ones, while unrounded back vowels are somewhat more central than rounded back vowels. Frequently, of the low vowels only the back member is ever rounded, and many languages avoid low round vowels altogether. The vowels of conservative Korean, given in (2), illustrate the avoidance of low round vowels. (In the modern varieties, there are seven vowels. The vowel [æ] has merged with [ɛ] for many speakers, while [y] and [ø] have become [wi] and [we] in the speech of most speakers.) Example words are given in (3).

<table>
<thead>
<tr>
<th></th>
<th>Front unrounded</th>
<th>Centralized front rounded</th>
<th>Centralized back unrounded</th>
<th>Back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>y</td>
<td>u</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>ɛ</td>
<td>o</td>
<td></td>
<td>o</td>
</tr>
<tr>
<td>Low</td>
<td>(æ)</td>
<td>a</td>
<td></td>
<td>ʌ</td>
</tr>
</tbody>
</table>

(3) | gi ‘era’ | gy ‘ear’ | gu ‘he’ | gu ‘sphere’ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>gæ</td>
<td>‘crab’</td>
<td>kʊ</td>
<td>‘craftiness’</td>
<td>no ‘paddle’</td>
</tr>
<tr>
<td>gæ</td>
<td>‘dog’</td>
<td>na</td>
<td>‘I’</td>
<td>nʊ ‘you’</td>
</tr>
</tbody>
</table>
While a central position of the tongue bunch is thus quite common, central vowels pattern like back vowels if they are unrounded, and like front vowels if they are rounded.

2.4.2 Diphthongs

When two different vowels appear in the same syllable, the combination is known as a **diphthong**. English has diphthongs in [lai] *lie*, [næʊ] *now* and [dʒɔɪ] *joy*, while German has the diphthongs [ai], [au] and [ɔi] in [tsait] *time*, [haus] *house* and [ɔɪç] *you (OBJ PL)*, respectively.

2.4.3 Nasalization

Vowels can be **nasalized**: if during their production the soft palate is lowered (so that the nasal cavity is opened up), a nasal quality is added to them. Nasalized vowels occur in Portuguese and French, and are symbolized by placing a tilde over the symbol. Thus, French has three nasalized vowels, occurring in [vɛ] *wine*, [tẫ] *aunt* and [ʒɔɪ] *you*.

**Q16 Are there any close, close-mid or open-mid front vowels in the English phrase *I’m not here to make friends?* Any velar sounds? Any fricatives whose friction is produced upstream from the larynx?**

The acoustic structure of vowels is determined by the resonances of the air in the vocal tract which modify the glottal waveform. These resonance frequencies or formants are counted from lowest to highest (formant 1, or F1, is thus the lowest). F1 corresponds to the degree of opening of the oral cavity. It is low for close vowels like [i y u] and high for an open vowel like [a]. F2 approximately depends on the distance between the lips and the main constriction in the oral cavity: a forward constriction as for [i] results in a high F2, while a velar or uvular constriction results in a low F2. The effect of F2 is similar to the resonance heard in a bottle under a tap: as it fills up towards the beginning of the neck and the column of air gets shorter, the resonance caused by the splashing water is higher. F1 is rather related to the size of the opening of the bottle: a thin neck creates a duller effect (low F1, as for close vowels) than a wide neck (high F1, as for open vowels). In the next chapter we will see how the distance between the lips and the constriction is a determining factor in explaining the frequencies of occurrence of vowel sounds in the languages of the world.

You can hear the effect of your second formant independently of the laryngeal source sound when you whistle. High notes are produced with the tongue in the position for [i] (or, better, [y], because of the lip rounding), low notes with the tongue in the position for [u] and intermediate notes with intermediate tongue positions for high vowels, like [u]. You can hear your first formant if you close your glottis (see Q13) and flick a finger against the side of your throat going from [i] via [e], [a], [o] to [u], once per vowel position. Figure 2.5 plots the first two formant frequencies of a male speaker of the seven Italian vowels in (1). The horizontal axis represents F2, with its origin on the right, while the vertical axis represents F1, with its origin at the
top. Both axes are non-linear, with higher values being more compressed than lower values. By arranging the axes in this way, the vowels get to be spatially arranged as in a conventional impressionistic vowel diagram or vowel trapezoid (see the vowel trapezoid in the IPA chart on page ix) or as in diagrams like those in (1). Incidentally, you may now see why people started calling the fundamental frequency the ‘f0’.

Because of the shorter length and smaller width of a child’s vocal tract, the formant frequencies of vowels produced by children are higher than those of adults, particularly those of F2. Likewise, women, whose vocal tracts are approximately 15 cm long as compared to 17.5 cm in men, have higher formant frequencies than men for the ‘same’ vowels. Listeners normalize for those differences after extracting the information about the sex and maturation of the speaker from the speech waveform.

2.5 CONSTRUCTIONS

Various types of constriction can be made in different locations in the vowel tract. We will first go through the different places, and then discuss the different kinds of constriction that exist for consonants. The expression ‘to articulate with X’ is used to mean ‘to form a constriction at X’.

Figure 2.5
Plots of F1 and F2 of seven Italian vowels as spoken by a male speaker in a space with inverted, logarithmic axes. By courtesy of Antti Iivonen, http://www.helsinki.fi/speechsciences/projects/vowelcharts/.
2.5.1 Places of articulation

Pharyngeal

The root of the tongue articulates with the back wall of the pharynx. Gulf Arabic has a voiceless pharyngeal fricative, as in [laḥam] ‘meat’.

Dorsal

The dorsum articulates with the roof of the mouth. If it is the back that articulates with the soft palate, the term velar is used. This place is used for [k] in French quand ‘when’, for [ŋ] in English hang, for [g] in English good and for [x] in Scottish English loch. If it is the front which articulates with the hard palate, the term palatal or palatoalveolar is used. Examples are [j] in Spanish yo ‘I’, [c] in Dutch ‘kacjɔ’ ‘cat+dım’ or [j] in English ship. These types of segments also involve a raising of the crown, and are therefore treated under ‘Coronal’ below. Fronted velars, like [ç] in German nicht or [k] in French qui, should be distinguished from palatales. Retraction of the dorsum allows the back of the tongue to articulate with the uvula. A voiceless uvular fricative [χ] occurs in Western Dutch, as in [χɛːl] ‘yellow’. The uvular stops [q g] occur in Tlingit, for instance.

Coronal

The crown may articulate with the upper teeth, the alveolar ridge or the forward part of the hard palate immediately behind the alveolar ridge. When the crown articulates with the upper teeth, dental segments are produced, like [θ] in English thing and [ð] in this and that; the dental plosives [t d] occur, for instance, in Sinhalese [tadə] ‘hard’. The label alveolar is used if the crown articulates with the alveolar ridge. Examples are [t] and [d] in German [tʰuːn] ‘do’, [duː] ‘you’, [n] in [nækt] ‘night’ or [l] as in [ˈɔləs] ‘everything’. If the crown articulates with the rear edge of the alveolar ridge, a postalveolar consonant is produced. English [ʃ] in shore, [ʒ] in measure, [ʧ] in tip and [dʒ] in jet are articulated with the crown of the tongue, while the front of the tongue is raised towards the hard palate. Often, as in Dutch, the tip is held behind the lower teeth for this type of consonant, in which case the contact is alternatively labelled prepalatal. A postalveolar articulation with just the tip of the tongue occurs in English [tʃ] as in try. If this type of contact is made with the tongue tip or the underside of the tongue blade with the tongue tip curled back, the term retroflex is used. Like many languages spoken in India, Hindi has retroflex consonants like [t d η], as in [tʰoŋdi] ‘cold’ and [gʰoŋ[a] ‘hour’.

Labial

If the lips articulate with each other, as in English [p b m] in spot, bell, mad, the place is bilabial. If the lower lip articulates with the upper teeth, the place is labiodental. It is used for [f] in German [fiː] ‘cattle’ or French [fɛʁ] ‘do’. 
2.5.2 Types of constriction

A first subdivision distinguishes between two kinds of constriction.

1. A constriction that is tight enough to lead to friction when a (voiced or voiceless) airstream is passed through it, as used with obstruents.
2. A type of constriction that allows a voiced airstream to pass through without friction, which is used for sonorants.

In the case of obstruents, an acoustic source is actually created at the point of articulation: either a popping sound is produced (for plosives), or friction is produced at that spot (for fricatives and affricated plosives). The auditory quality of sonorants relies exclusively on the different shapes the vocal tract is given, i.e. on the resulting modifications of the acoustic characteristics of the sound produced by phonation in the larynx.

Obstruents are subdivided into plosives (also called stops), fricatives and affricates.

Plosives

These are formed by creating a complete closure at some point in the speech tract, behind which the air from the lungs is compressed until the closure is abruptly released so that the air explodes outwards. Since the soft palate is raised, the air cannot escape through the nasal cavity. Examples are French voiceless [p t k], as in [pip] 'pipe', [tip] 'type', [ekip] 'crew'. Voiced plosives occur in French [bide] 'bidet' and [gã] 'glove'. Plosives have a very brief friction burst when they are released, which is not usually heard as friction but is responsible for the popping quality of plosive releases.

Fricatives

These are formed by narrowing the speech tract to such a degree that audible friction is produced when air passes through. English has the voiceless labiodental fricative [f] infee, the voiceless dental fricative [θ] in thigh, the voiceless alveolar [s] in sigh and the voiceless palatoalveolar [ʃ] in shy. The voiced counterparts [v ð z ʒ] occur invie, that, zoo and measure, respectively. (At the beginning of the syllable, English [v ð] are frequently pronounced without friction.)

Affricates

Affricates are plosives whose release is slow instead of sudden, causing a longer phase of turbulence. The affricates [pf] and [ts] occur in German ['pfaifa] 'pipe' and [tsait] 'time', and the palatoalveolar affricates [tf] and [dʒ] occur in English cheer and jeer, respectively.

Sonorants

These divide into nasals and approximants.
Nasals

For nasals the soft palate is lowered, and the oral cavity is blocked completely at some point. A slow, deliberate pronunciation of morning will allow us to observe how each of the three nasal consonants in the word has a different place of articulation: a bilabial [m], an alveolar [n] and a dorsal [ŋ]. (Pre-)palatal [ɲ] occurs in Dutch, as in ['spaɲɔ] ‘Spain’, and in French, as in [aɲ] ‘lamb’.

Approximants

Approximants derive their name from the approximation of the articulators, which gives rise to a light or near-contact, the airstream being so weak that no friction is produced. The (pre-)palatal approximant [j] occurs in English yes, while a bilabial one ([β]) occurs in Southern Dutch [βæt] ‘what’. (When combined with a symbol for a voiced fricative, the subscript [,] indicates an approximant, i.e. frictionless pronunciation.) Frequently, languages have rhotics, or [r]-type segments, that are approximant, such as the palatal replacement [j] in English ray. Also, trills occur, during which the uvula is allowed to vibrate against the back of the tongue [r], as in French [pari] ‘Paris’, or the tongue tip against the alveolar ridge [ɾ], as in Spanish [ˈpero] ‘dog’. When, instead of a series of vibratory taps, a single brief contact is made, a flap is produced ([ɾ]). In Spanish and Catalan the alveolar flap contrasts with a trill, a minimal pair in Catalan being [ˈparo] ‘father’ – [ˈpara] ‘grapevine’. In Arawak, the alveolar flap contrasts with a retroflex flap, which involves flicking the curled-back tongue forward, causing the tip to hit the rear edge of the alveolar ridge, as illustrated in [ˈhɔɾo] ‘swampy’ – [ˈhɔɾɔɾo] ‘cloud’ (Pet 1979), while Toda contrasts three places of articulation for trills (or flaps), postdental, alveolar and retroflex, as in [kəɾ] ‘border of cloth’, [kar] ‘juice’ and [kaɾ] ‘pen for calves’ (Spajić et al. 1996).

For [l], as in German [ˈalas] ‘everything’, the airstream is partly blocked by the tongue tip contact with the alveolar ridge, but allowed to escape freely on one or both sides. Sounds which have this type of partial occlusion are called lateral. Because the air escapes without friction through the lateral opening, the German lateral is an approximant. The voiceless lateral fricative ([l]) has a turbulent escape of air along the lateral opening(s); it occurs in Welsh, as in the place name [lan’didno] (Llandudno). Laterals are usually alveolar: the crown articulates with the alveolar ridge. A palatal (approximant) lateral [ʎ] occurs in Italian [ˈzbaʎːo] ‘mistake’ or Catalan [ˈkɔp] ‘wolf’.

2.6 SEGMENTAL DURATION

In many languages, duration contrasts exist in the group of vowels, as in Czech and Hawaiian, or in the group of consonants, as in Italian and Tamazight, or in both, as in Finnish and Japanese. Long consonants are also called geminates. Italian and Japanese have long consonants in intervocalic position, while other languages may additionally have them in word-initial and word-final positions, like Swiss German
(Kraehenmann 2001). An illustration of such a quantity contrast is Italian [fato] ‘fate’ versus [fat:o] ‘fact’. Frequently, only a subset of the consonants occurs both long and short. Thus, only [p t k s] occur as long consonants in the native vocabulary of Japanese. Vocalic quantity contrasts, such that exist in Hawaiian, are more common. As in the case of consonantal duration contrasts, sometimes only a subset of the vowels occurs both long and short. Thus, Chipewyan has five short vowels ([i e a o u]), but only three of them have long variants ([i: a: u:]). Moreover, the quality of the long vowels may differ somewhat from that of the short vowels. Finnish, for instance, has eight monophthongal vowels appearing both short and long, but long [e:] has a closer quality than short [e].

In German, the durational contrast coincides with a difference in the position of the root of the tongue, which is advanced somewhat during the pronunciation of the long vowels. Such vowels are known as tense, while their unmodified counterparts are lax. As a result of the widening of the pharynx, the tongue body will tend to be higher for tense vowels than for lax vowels, at least for vowels that are not fully open. In (4), which gives the vowel system of Standard German, the symbol on the left in each box is a lax, short vowel, while the symbol on the right represents a long, tense vowel. Examples of words with these vowels can be found on page 81.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Front unrounded} & \text{Centralized front rounded} & \text{Centralized back unrounded} & \text{Back rounded} \\
\hline
\text{High} & i:i & y:y & o:u: \\
\text{Mid} & e:e & æ:ø: & o:ø: \\
\text{Low} & æ: & a:a & \\
\hline
\end{array}
\]

In more conservative German, there is also a long lax vowel [æ:] (bracketed in (4)), which has merged with [e:] in the speech of many speakers.

### 2.7 Complex Consonants

Complex consonants are single segments which in some way have two distinguishable articulations. We distinguish consonants with a secondary articulation, consonants with a double articulation, and manner-contour consonants.

#### 2.7.1 Secondary Articulations

The articulation of a consonant does not require the services of the entire tongue as well as the lips. For [t d], for instance, only the crown and the sides of the tongue are used, and for [f p m] only the lips. In fact, the lips and tongue can be used to form a vocalic articulation simultaneously with the production of the consonant. The result is what is known as a consonant with secondary articulation. The following types of secondary articulation can be distinguished:

1. **Labialization.** During the articulation of the consonant, the lips are rounded. A labialized velar plosive, for instance, is symbolized [kʷ].
2 **Palatalization.** The front of the tongue is raised (as for [i] or [j]) during the pronunciation of a consonant. A palatalized bilabial plosive, for instance, is symbolized [p’].

3 **Velarization.** The back of the tongue is raised (as for [u] or [ɯ]) during the pronunciation of a consonant. A velarized alveolar lateral approximant (‘dark l’), symbolized [ɬ], is used postvocally in most varieties of English, as in *all.*

4 **Pharyngealization.** The root of the tongue is retracted towards the back wall of the pharynx. Arabic ‘emphatic’ consonants are pharyngealized. This is indicated by a superscript voiced pharyngeal fricative symbol after the consonant symbol (e.g. [tˤ]) or by a dot placed below the phonetic symbol concerned (e.g. [ṣ]).

### 2.7.2 Double articulations

Some consonants have two consonant-type constrictions at the same time, such that it is not possible to say which is ‘primary’ (i.e. consonant-like) and which ‘secondary’ (vowel-like). Well-known examples are [kp] and [gb], labial-velar plosives, which occur in many Niger-Congo languages. The English consonant [w] is a labial-velar approximant: it is pronounced with a raising of the back, as well as with rounding of the lips. Such consonants are said to involve a **double articulation.**

### 2.7.3 Manner-contour consonants

Some consonants change their constriction-type half-way through. An example is provided by **prenasalized stops,** which occur for instance in Bantu languages. Such consonants begin like nasals and end like plosives at the same place of articulation ([mb], [nd], [ŋg]).

### 2.8 NONPULMONIC CONSONANTS

We briefly describe the three classes of nonpulmonic segments here.

1 **Clicks.** Clicks are produced with the help of a velar closure (as for [k]) plus a closure somewhere further forward. This forward closure may be located at the lips (rare) or at, or immediately behind, the alveolar ridge. The trapped air in the pocket in front of the velar closure is rarefied by lowering the body of the tongue or, in the case of a bilabial click, the jaw. When the forward closure is released to allow outside air to rush into the pocket of rarefied air, a clacking noise burst results. This way of creating an air pressure difference is known as the **velaric airstream mechanism** (Abercrombie 1967). Clicks may be contrastively accompanied by a glottal stop at the beginning of the postclick vowel, by nasalization on the following vowel, by aspiration preceding the following vowel or by a slow (affricated) release of the velar closure. Five articulation places for the forward contact occur. Their
symbols are included under ‘Other symbols’ in the IPA chart on page ix. The alveolar click is used paralinguistically in English as a sign of disapproval (rendered *tut, tut* in British English and *tsk, tsk* in American English), while repeated lateral alveolar clicks are sometimes used to imitate the noise a horse’s hooves make on the pavement. Clicks only occur in languages spoken in southern Africa, !Xù, Nama and Xhosa being well-known examples of such languages.

2 **Implosives.** Implosives have a closure as for plosives, as well as a closed glottis. By lowering the larynx, the air in the mouth and pharynx is rarefied, so that a noise burst will result when the oral closure is released. Because, on its way down, the speaker relaxes his or her glottal closure, phonation will occur as the glottis meets the air in the trachea. Such speech segments occur in Vietnamese, as well as in some Niger-Congo languages. Their symbols are [ɓ ɗ ɠ].

3 **Ejectives.** The initial articulatory configuration of ejectives is like that used for implosives, but instead of being pulled down, the larynx is pushed up. The air inside the pharynx and mouth is compressed, so that upon the release of the oral closure, an egressive noise burst occurs. Ejectives are reasonably common and are symbolized [p’ t’ k’]. Implosives and ejectives are produced with the glottalic airstream mechanism.

### 2.9 Stress

In many languages, words consist of rhythmic units called **feet**, the most common foot type being disyllabic. One of the syllables of the foot is more prominent or stronger than the other syllable(s) in it, which for this reason is called the **stressed** syllable. In city, this is the first syllable. A word like celebration contains the two feet [ˌsɛlə] and [ˌbreɪn]. One of the feet in a word has the **primary stress** or **word stress**. In celebration, this is the last foot. The other feet have **secondary stress**. The IPA notation for primary stress is [ˈ], and for secondary stress it is [ˌ], to be placed before the syllables concerned. The stress pattern of celebration contrasts with that of [ˈæləˌgɪtə] alligator in the location of the primary stress. In addition to disyllabic feet, there may also be monosyllabic feet, as in cat, and ternary feet, as in origin [ˈɔrədʒɪn]. In many publications on this topic, [ˈ] is used over the vowel for primary stress, and [ˌ] for secondary stress. In this book these accent marks are used only for tone (section 2.2.4 and chapter 10).

### 2.10 Conclusion

This chapter has outlined the workings of the speech production mechanism in a way that will enable you to follow the discussion in the rest of this book. For many users of the book, it will have served as a brief refresher course, while for others, who may be new to the topic, it will have served as an introduction.
NOTE

1. There is an additional valve, called the epiglottis, positioned above the larynx where the root of the tongue begins. It normally points upwards, but it flaps down to channel food and saliva into the oesophagus – the tube behind the larynx leading to the stomach – when we swallow.
The phonologies of different languages are in many respects very similar, to the extent that some features appear to be part of every language. These cross-linguistic similarities are due to two sources. One is the structure of the human brain. Even though we don’t understand the workings of this complex neurological organ, it seems obvious that the physiological basis for the hierarchical structure of language is to be located in the way the brain operates. Our awareness of a linguistic expression like *Many pens leak* will concern simultaneous structures at different levels in each of two structural hierarchies: it is a sentence, and there are two phrases, three words, four syllables, a string of 11 segments, and a prosodic structure, which will be similarly hierarchical. When viewed as structural layers, linguistic hierarchies must be processed for production and perception purposes with parallel clock times at different frequencies, one for segmental articulations, one for stress, and so on (Poeppel 2003). Other aspects are more clearly due to the ergonomics of the speech process. Language users, and thus languages, prefer distinctions that are easy to perceive and easy to produce. The difference between [t] and [n] is very clear to the perceiver and not too difficult to make for the producer. The contact made by the tongue tip and rims with the upper gums is the same, while the velum is lowered for [n], opening up the nose at the back, and raised for [t], trapping the air behind the oral closure. In the open-velum position, very little effort is needed to bring the vocal folds somewhat closer together than during breathing in order for them to start vibrating. This is so easy that voicing during sonorant consonants and vowels has been called ‘spontaneous vibration.’ By contrast, if we block the egressive flow by closing off both the mouth and the nose, it will take some effort to create a sufficiently powerful air stream through the glottis to make them vibrate at all: we are pumping more and more air into a small, closed pocket of air. Speakers would be well advised therefore not to be too eager to vibrate their vocal folds while their vocal tract is significantly obstructed, because maintaining an air pressure difference across the glottis takes some effort. (You can increase the pressure from the lungs, or make more room by pulling up the velum for [d ɣ] or blowing out your cheeks for [b].) That is why almost all languages have a dental/ alveolar [t] and [n], but only 64% have the voiced counterpart of [t] ([d]), and fewer than 0.5% the voiceless counterpart of [n] ([n̥]). It would be extremely improbable to find a language that had no [t n] but did have [d ɣ]. Another speaker interest is to duplicate contrasts. If a language has the vowels [i e a o u] and you find it has a nasalized [õ], it would be a good bet to assume it also has [ã], while the presence of [ē] is almost as
likely. The speaker’s phonetic routine of nasalization is exploited so as to maximize contrasts with the same velum-lowering gesture. But here too the hearer is not forgotten. For high vowels like [i u], nasalization does not have a whole lot of acoustic effect, and many languages, like French, therefore leave them out of their subsystem of nasal vowels (‘system gaps’).

Low-cost contrasts are thus frequent. Some sound contrasts are so easy to make and so clear to the ear that they are found in every language that has been described. For instance, no language has been found without vowels, and no language has been found without consonants. There are two responses to this state of affairs. One is to assume that whatever languages have in common is universal and that the explanation for the universality is that this is hard-wired in our brains, innate. The other is to assume that the neurological, physiological, physical and social conditions under which languages arose and developed are the same across our species and the forces that determine their structure are the same. Those forces must allow for a fair number of degrees of freedom in order to explain the variation that is seen. Apparently, after going for the obvious sound contrasts, which occur in the great majority of languages, weighing up the cost to the speaker and the benefit for the hearer may lead to a large number of options. Also, there are apparently other factors that explain why some languages have many sound contrasts, and others few. The only factor that we can at this point be certain of is the historical dimension: the phonological complexity of languages changes very little in its transfer from one generation to the next. Old English, for example, had broadly the same level of complexity as most contemporary varieties of English, even though it has changed virtually beyond recognition after some 1,500 years of development.

Q17 Why is [i e a ə o ŋ u] an implausible vowel system?

Q18 Assume that all the languages that were ever spoken in the world have at least three vowels. What explanation would be given of this fact by someone who rejects the theory that human brains are genetically programmed to have at least three vowels?

3.2 VARYING COMPLEXITY

Languages differ in their morphosyntax just as they differ in their phonologies. To begin with, they differ greatly in the number of segments they have. In UPSID (see note 1), a corpus of segment inventories of 451 languages (approximately 7% of all the languages of the world), the smallest number appears to be 11 (e.g. Rotokas, spoken in Papua New Guinea) and the largest a staggering 141 (!Xū, spoken in Namibia and Angola) (Maddieson 1984: 9). And when two languages have the
same number of segments, they are unlikely to have identical sets. Another language that, like Rotokas, has 11 segments is Pirahã, which is spoken in Colombia (Everett 1986). It shares [p t k g j i o a] with Rotokas, but while Rotokas has [e u] and the consonants [f] and a flap, Pirahã has [f b s h].

Second, different languages will have different constraints on the way segments are combined to form syllables. The initial consonant(s) of the syllable are known as the **onset**, the vowel is in the **peak**, and the closing consonant(s) form the ** CODA**. Coda and peak form a constituent called the **rhyme** (also spelled * rime*). In (1), we show these constituents in a tree diagram.

\[\text{(1)}\]

\[
\begin{array}{c}
\text{ONSET} \\
\text{C-} \\
\text{PEAK} \\
\text{V} \\
\text{CODA} \\
\text{-C}
\end{array}
\]

According to Blevins (1995), the lowest degree of complexity in syllable structure is represented by languages that have a single (short) vowel in the peak and allow maximally one consonant in the onset. The syllable structure of such languages is (C)V. Further complexity can be achieved in a number of ways:

1. The onset may be obligatory. Languages with obligatory onsets are not hard to find, like Arabic, Dyirbal and Klamath, but languages that allow only CV, i.e. an obligatory onset preceding an obligatory monomoraic vowel, must be rare, to the extent that we have no example. Analytical issues may arise from phonetic onsets in onsetless syllables. Many languages, like German, regularly have a glottal stop in onsetless syllables, as in [ˈɛfə], pronounced [ʔafə], ‘monkey’, while in Mba and Ndunga word-initial high vowels appear to be variably preceded by [ʔ], [fi] or a homorganic semivowel (Pasch 1986: 32, 91). Henan Mandarin has a homorganic semivowel preceding a non-low onsetless vowel, as in [ǐɪ⁴¹], pronounced [ji], ‘one, numeral’; [y⁴¹], pronounced [u], ‘fish’; [u⁴¹], pronounced [wu], ‘house’; [x⁴¹], pronounced [ux], ‘hungry’; and a glottal stop before low vowels, as in [æe⁴¹], pronounced [ʔae], ‘love’ (personal field notes, CG, 2014). Phonemic glottal stops are easily diagnosed in languages like Hawaiian, where they contrast with empty onsets, as in [aa] ‘jaw’ and [ʔaa] ‘fiery’ (Elbert and Pukui 1979), or Arabic, which has the same contrast in the coda of the syllable and has the glottal stop in the onset within words, as in [lawʔa] ‘sorrow’ (Thelwall and Sa’adeddjin 1990). German has none of these three properties.

2. There may be an optional coda. Languages that don’t allow codas include Hawaiian, Bakwiri and Fijian. If the language allows a coda, a further option is that the coda may be complex, i.e. may be a **cluster** containing more than one C. Spanish, Japanese and Italian allow only one C, but Klamath and French allow two.
Some languages allow the syllable peak, i.e. the structural position indicated by ‘V’, to be filled with a consonant. Thus, in Czech, the liquids [r l] appear in the peak, as in the geographical names [ˈbr̩.no], [ˈvɪ.tɑ.ʋa], Brno, Vltava, and American English has words like girl [grːl] and mountain [ˈmaʊn.ˈtɛ̃]. Zware Tamazight allows all consonants to be syllabic, even voiceless ones, as shown by [ɪ.ʒn.ˈmo.ʃn] ‘raisins’, [a.ˈdf.ˈfu] ‘apple’. In the second word, the nucleus (‘V’) of the stressed syllable is filled by a voiceless fricative (Gussenhoven 2017). Where languages agree in the number of consonants they allow in the coda or onset, they will still differ in how these consonant positions can be filled. In German, [kn- ʃn-] are possible onsets, as in [ˈkna.ː.bo] ‘boy’, [ʃn.ɛ.ka] ‘snail’, and so are [ps- ks-], as in [psy.ˈko.ˈlo.ɡj] ‘psychological’, [ˈksan.ˈtɛ̃] ‘Xanten’, all of which are lacking in English. (German [ts-] as in [tsaːr] ‘czar’ is usually interpreted as a single segment, an affricate.) Conversely, English has [l] in lure, which is absent in German. An unexceptional implicational relationship is that the presence of a complex constituent implies the permissability of less complex constituents. Thus, a language that allows two consonants in the onset will also allow one consonant in the onset.

Third, different languages will have different phonological processes. A process that occurs in British English but not in German is preglossation, by which [p t k] are [ʔp ʔt ʔk] when they occur at the end of a syllable before another consonant, as in ripped, mats, thickness. And an example of a process in German that does not exist in English is final devoicing, by which all obstruents (i.e. plosives and fricatives) are voiceless when they occur at the end of a syllable. This rule is responsible for the fact that in German no voiced fricative or plosive appears at the end of a word, even though voiced obstruents occur in the inflected forms of such words. For instance, the nominative form for ‘dog’ is [hont], while the genitive form is [hon.ˈdɒs]. Many languages have a rule of final devoicing, including Polish, Dutch and Catalan.

Q19 The following words, taken from Huisman et al. (1981), illustrate the syllable structure of Angaatiha (the period separates the syllables):
Some typology: sameness and difference

| kə.mə.ai       | ‘him’          |
| a.tiʔəɾə       | ‘thunder’      |
| ai.n.taʔo      | ‘bird type’    |
| ma.nji.njaiʔo   | ‘children (Objective)’ |
| ta.m.pwaiʔo    | ‘lizard type’  |

1. Give the AngaatiHa syllable structure as a CV formula, placing optional elements in brackets.
2. Does the language have syllabic consonants?

Q20

1. Can you think of (a) a segment which exists in English but not in your own language, and (b) a segment that exists in your own language but not in English? If your native language is English, answer the question for any foreign language you are familiar with.
2. What would the final consonant of the German word for Kiev be? Why?

3.3 UNIVERSALS AND IMPLICATIONAL RELATIONS

While it is clear that different languages may have very different phonologies, at the same time it is clear they have many things in common. For a start, all languages would appear to have syllables, and all segment inventories can be split into consonants and vowels. All consonant inventories include voiceless plosives; i.e. all languages have at least two of the three consonants [p t k]. Then, there are near-universals. For instance, only two languages in UPSID, Rotokas and Pirahã, have no sonorant consonants. Or, again, all languages in UPSID except Hawaiian (i.e. 99%) have some kind of [t], and 90% have [i]. It is also striking that the coronal place of articulation is much commoner generally, and also shows more subdivisions, than either the labial or the dorsal places of articulation.

When the group of more common segments is compared with the group of unusual segments, there are two observations to be made:

1. Unusual segments tend to occur in larger segment inventories. For instance, an unusual segment like [kʰ], a [k] with rounding and aspiration, typically occurs in languages with large consonant inventories, like that of Igbo, which has 20 pulmonic egressive plosives in addition to three implosives, or Haida, which has 46 consonants in all.
2. Unusual segments tend to be phonologically more complex than common segments. For example, a common segment like [k] (99.4%) just involves a complete closure between the back of the tongue and the soft palate. The
articulation of [k] allows organs of speech other than the back of the tongue to take the line of least resistance, requiring no accompanying actions of the vocal folds (like aspiration, or voicing during the closure, or a glottal closure), of the lips (lip rounding) or of the front of the tongue (palatalization). On the other hand, an unusual segment like [kʰ] requires the same oral closure as [k], but in addition has aspiration as well as lip rounding. The relationship is not absolute, however. Although [θ], for instance, is phonologically simple, it is nevertheless a rare segment.

What do these facts suggest about the phonological structure of language? The first fact suggests that languages 'build up' their phonologies in an ordered fashion. It suggests that a language will have segment X only if it already has segment Y. There are in fact many such implicational relationships, as Roman Jakobson had taught the world in 1928 (Jakobson 1990: chs. 9, 10, 18). Thus, no language has a voiceless nasal without also having its voiced counterpart, and no language has [z] without also having [s]. With very few exceptions, languages do not have front rounded vowels ([y ø]) if they do not also have front unrounded ([i e]) and back rounded ([u o]) vowels. As will be clear, most of these implications are tendential, and true only in a statistical sense. For instance, the presence of [f] generally implies the presence of [p], but Chuave is an example of a language for which this is not true. Apparently, there is no such thing as an absolute order in which languages avail themselves of the universal phonological resources.

The second fact suggests that one way in which languages construct their segment inventories is by adding elements to already existing segments. If we continue to use the metaphor of 'building up' the segment inventory, it is as if you begin by making some choices from a collection of fairly run-of-the-mill segments, and then, as you require more of them, select further elements which you can use to create more segments. For example, many languages have [p t k], while a smaller number have [p t k] and [b d g]. These languages can be seen to have increased their inventory by adding the element 'vibrating vocal cords' to the plosive segments they already had. Then, a smaller number still has these, plus [pʰ tʰ kʰ], for instance. Here, the language can be seen to have added 'aspiration' to the [p t k] it already had. (But, again, instead of 'aspiration', languages may employ other elements that serve to expand the segment inventory [p t k b d g].)

Q21 The mean number of consonants per language in UPSID is 22.8, with a range of 6–95. The mean number of vowels is 8.7, with a range of 3–46. In what respects are the following languages atypical?

1. !Xù has 95 consonants and 46 vowels.
2. Pawaia has 10 consonants and 12 vowels.
3. Haida has 46 consonants and 3 vowels.
4. Norwegian has 22 consonants and 19 vowels.
As Jakobson (1968) emphasized in a famous monograph, children learn the sounds of their language in a particular order, regardless of the language, which reflects the scale of complexity implied in this paragraph: the unusual tends to be rare and tends to be acquired late. Thus, the metaphor of ‘building up’ the phonology in a sense comes true every time a child learns a language.

3.3.1 Plain or special?

A third observation is made by Maddieson (1984): the number of vowels and the number of consonants are positively correlated. That is, when one language has more consonants than another, it is likely that it also has more vowels. This suggests that languages that are more complex in one area of the phonology (in this case, the consonant inventory) also tend to be complex in other areas (in this case, the vowel inventory). It is apparently not the case that complexity in one area is somehow compensated for by simplicity in other areas: phonologies differ in complexity. For instance, languages with larger inventories tend to have more complex syllable structures. That is, languages that allow only simple syllable structures like CV tend to have smaller inventories than languages that, like English and German, allow up to three consonants at the beginning of the syllable (cf. English [strə] *straw*, German [ʃtʁo:] ‘straw’), as well as more than one at the end (cf. English [læmp] *lamp*). In tone languages, the number of different pitch patterns typically lies between two and six, but the number of such patterns does not appear to be greater in languages with smaller inventories or in languages with simpler syllable structures. UPSID does not list permitted syllable structures, or prosodic distinctions, so the facts here are not easy to give. Maddieson (1984: 22) does, however, present the total number of possible syllables in a small corpus of languages. The total number is determined by (a) the number of different vowels and consonants, (b) the permitted syllable structures and (c) the number of tonal (prosodic) distinctions. If complexity in one area of the phonology goes hand in hand with complexity in another, then the greater the number of possible syllables in a language, the larger should be the number of segments, and the larger should be the number of permitted syllable structures, as well as the number of prosodic distinctions. In the representative list of languages Maddieson gives, this is indeed the case. To give some idea of the multiplicative effect of these three factors: the language with the largest number of different syllables in Maddieson’s minicorpus (Thai) has 146 times as many possible syllables as the language with the smallest number (Hawaiian).

Q22 In a corpus of 10 languages, the number of segmental contrasts was found to be related to the mean word length expressed as the number of segments in the word, in the sense that the smaller the segment inventory was, the greater the mean word length was (Nettle 1995). Speculate on the cause of this negative correlation.
### 3.3.2 Avoiding complexity

There is another, somewhat more tenuous, relation worth drawing attention to. Unusual segments are not unusual just because they occur in relatively few languages; they also tend to be less frequent in the languages that have them. That is, as languages increase their phonological resources, they tend to do less with them. For instance, Dutch has front unrounded, back rounded and front rounded vowels. First, the set of front unrounded vowels contains one segment more than either the front rounded or the back rounded sets. Then, in almost every case, the frequency of occurrence of the front rounded vowels is lower than those of the corresponding front unrounded or back rounded vowels. The figures in Table 3.1 are based on the frequencies of the vowels in the 1,000 most common Dutch words, expressed as percentages of the total number of vowels. The frequency of occurrence of these vowels amounts to 20% of all the segments. What can also be seen by comparing the figures in the second row with those of the first and fourth rows is that lax (short) vowels are more common than tense vowels.

### 3.3.3 A word of caution

You may by now have gained the erroneous impression that it is in fact easy to count the segments, or syllables, or prosodic distinctions, in a language. We hasten to dispel this notion: it is not easy to count these things at all. There are two reasons for this. One is that languages frequently have what are sometimes called marginal segments or patterns of arrangements (Moulton 1962). These are restricted to onomatopoeic words, in which the phonology echoes the meaning of the word, and (recent) loanwords. For instance, Dutch has a number of such vowels, among which [ɛː], which occurs in loans (e.g. [krɛ:m] ‘cream’) and in the onomatopoeic [ˈkleːrə] ‘cry, bawl’. Should it be counted as a Dutch vowel? Or, again, no Dutch words begin with [fj]-, except [fjɔrt] ‘fjord’. Is [fj-] a Dutch onset, or is it a Norwegian onset which speakers of Dutch have taken in stride? Depending on the answers to such questions, the number of Dutch vowels will vary between 16 and 25. The second reason why it is difficult to count these things is that the count will depend on the analysis. Suppose a language has the five vowel qualities [i e a o u]. Suppose further that they can be either long or short and, moreover,

---

**Table 3.1**

Proportional frequencies of occurrence (%) within the group of Dutch vowels of front unrounded, front rounded and back rounded vowels (adapted from van den Broecke 1976).

<table>
<thead>
<tr>
<th></th>
<th>Front unrounded</th>
<th>Front rounded</th>
<th>Back rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tense, short</td>
<td>i</td>
<td>y</td>
<td>u</td>
</tr>
<tr>
<td>Lax, short</td>
<td>13.53</td>
<td>2.33</td>
<td>12.12</td>
</tr>
<tr>
<td>Lax, short</td>
<td>22.02</td>
<td>0.29</td>
<td>10.23</td>
</tr>
<tr>
<td>Tense, long</td>
<td>12.32</td>
<td>0.29</td>
<td>10.23</td>
</tr>
<tr>
<td>Diphthong</td>
<td>9.36</td>
<td>1.79</td>
<td>1.70</td>
</tr>
<tr>
<td>Total</td>
<td>66.93</td>
<td>6.30</td>
<td>26.77</td>
</tr>
</tbody>
</table>

---
that every combination occurs as a diphthong. The number of different vowels in a sense is 5, but in another sense it is 30. UPSID would in a case like this give the number as 5, but if, say, the combinations [ae], [uu], [eo] had been missing from the language, then, paradoxically, the number of vowels would have been given as 27. Because of the large number of languages involved, the trends noted in UPSID are unlikely to depend very much on how these decisions are arrived at. What is important is that you should not get the idea that it is easy to say what a segment is, and that counting things is no problem.

3.3.4 Speech ergonomics

Clearly, languages somehow monitor the development of their phonologies, and check segments and inventories off against two very general guidelines: ‘Don’t make things difficult for the speaker’ and ‘Don’t make things difficult for the listener’. That is, the best systems are those in which contrasts are maximally distinct with the least amount of articulatory effort (Flemming 1995; Boersma 1998). The reasons for specific statistical tendencies may therefore be either articulatory or perceptual. In some cases, the speaker’s and the listener’s interests may go hand in hand, but often the explanation of the statistical fact lies in either the speaker’s interest or the listener’s:

1 Plosives are more common than fricatives. Plosives require a brief closure of the oral tract in order for air pressure to build up behind it. This is all it takes to create a brief friction burst at the release. By contrast, fricatives require considerable airflow for the full duration of the consonant to keep the air turbulence going and produce audible friction. The speaker’s interest here favours plosives over fricatives.

2 Voiceless plosives are more common than voiced ones. With every opening action of the vibrating vocal folds, the speaker releases some air from the trachea into the vocal tract. Because the vocal tract is closed off during the closure phase of the plosive, the air pressure in it will rise until it equals the air pressure below the vocal folds that drives their vibration. Just as we stop pushing the lever of a bicycle pump when the tube is full, so we will avoid a continuation of the vocal-fold vibration during the closure stage of the plosive. No such ergonomic conflict arises during voiceless plosives, there being no requirement for any airflow during an open glottis. Again, the speaker’s interest is at stake here.

3 Voiceless fricatives are more common than voiced fricatives. The relatively low airflow that results from the release of a rapid series of small air puffs into the vocal tract during vocal-fold vibration conflicts with the requirement of a sufficiently generous airflow through the narrowed passage where the fricative is articulated. As a result, voiced friction is weaker than voiceless friction. Arguably, since voiceless friction is perceptually more distinct from the acoustic properties of sonorant consonants and voiced plosives, speaker and hearer interests go hand in hand here.
Front rounded and back unrounded vowels are less common than either front unrounded or back rounded vowels. When you place an empty bottle under a running tap, the resonance in the bottle set up by the jet of air hitting the bottom of the bottle or the surface of the water already in it increases from low pitched to high pitched. Starting from [u], you can reproduce this effect by moving into the articulatory position for [i]. At [u], the lips are rounded, pouted even, so as to increase the distance between the raised back of the tongue and the aperture at the lips, creating the longest possible distance between them. But when saying [i], we spread the lips, so as to shorten the distance between the raised front of the tongue and the lip aperture. While the articulation of [i] corresponds to the fullest state of the bottle and that of [u] to the emptiest one, front rounded [y] and back unrounded [ui] represent intermediate positions, which are less likely to be used than the perceptually more extreme vowels. Here, the hearer’s interest is the decisive factor.

Ease of production is most clearly seen in the tendency for particular articulations to persist. Prenasalized stops, for instance, always have the same place of articulation for the nasal and the oral stop, as in labial, coronal and dorsal [mb nd ng]: segments like *[mk] or *[nb] have not been attested. Similarly, nasal consonants in the coda tend to share the place of articulation of the following plosive, as in Japanese [ram.pu] ‘lamp’, [jon.da] ‘read’ and [man.gan] ‘manganese’. And most languages have rules of assimilation, which cause some articulatory feature of one segment to be transferred to an adjacent one, such as when English in is pronounced [im] in in Paris or [in] in in Copenhagen. While for the speaker there are obvious advantages in extending the scope of an articulatory gesture, for the listener it is better to be able to hear differences. Paradoxically, therefore, there is also the tendency for languages to avoid repetition of the same thing. For instance, many languages that have labialized consonants lack labialized labials. Thus, Bakairi has [t ara k ara d ara g ara] but lacks [p ara b ara], even though the nonlabialized plosive series shows no gaps (Wetzels 1997a). Likewise, there are many languages that allow complex onsets like [pn- kn-] or [pl- kl-] but disallow [tn-] and [tl-], which last combinations have a single place of articulation. In chapter 9, we will introduce and exemplify the Obligatory Contour Principle, which is held to be responsible for this avoidance of repetition.

### 3.3.5 System gaps

We have seen that the segment inventories of languages tend to be constructed as if languages drew on the stock of phonological resources by adding elements to sets of segments. By adding to [p t k] the element ‘vocal-fold vibration’, you will produce the series [b d g]; further adding ‘aspiration’ will produce [p b t k], as in Burmese, for instance, and so on. Or if you have five vowels, you may allow them to be long as well as short, or you may nasalize them, or provide them with pharyngealization. (This is part of the answer to the secret of the 46 vowels of !Xu.) If
phonological resources are typically made available per set of segments rather than per individual segment, as indeed shown by Clements (2004), you may well wonder why languages so often have system gaps. For instance, Dutch has the voiceless unaspirated series [p t k] but only the voiced [b d]. So where is [g]? Dutch is not alone in having this gap. It is in fact the most commonly occurring gap among the voiced stops (e.g. Czech, Hixkaryana, Thai). And [p] is the least common among the voiceless plosives, being absent from languages as diverse as Arabic, Chuave, Dizi, Hausa, Vietnamese and Yoruba. In the examples given here, there appears to be a relation between the gaps and efficiency. The voiceless plosive [p] is relatively inefficient from the point of view of the listener, because the stop burst, which is one of the major cues to the presence of a plosive, is of much lower intensity in the case of [p] than in the case of other plosives, due to the lack of a resonating cavity in front of the point of release, where the burst is created (Stevens 1997: 494). The voiced [g] is relatively inefficient from the point of view of the speaker, because the relatively small air cavity behind the velar closure causes the air to accumulate fast below it, increasing the supraglottal air pressure and diminishing the glottal airflow, thereby causing voicing to stop (Ohala 1989). That is, [p] is relatively hard to hear, and [g] is relatively hard to say. Again, while these statistical tendencies are very clear, languages may deviate from them. Thus, Hawaiian, quite exceptionally, has [p k] but not [t]. In this case, an earlier stage of the language did have [p t k], but somehow [t] was replaced with [k], after [k] had become [ʔ].

Q23 The Zaza variety of Kurdish has the voiceless plosive series [p t c k q], but in the voiced series there is one missing consonant (Todd 2002). Which one would you guess this is?

Q24

1 On average, voiceless plosives are more frequent than voiced plosives, and coronal segments are more frequent than noncoronal segments. This holds true both for the occurrence of these segments in the inventories of the languages of the world, and for their occurrence in the words (or texts) of any individual language. To represent this situation graphically, draw a set of two coordinates, with frequency of occurrence on the y-axis and the three places of articulation labial, coronal and dorsal on the x-axis. Draw two theoretical graphs, one for the voiceless plosives and one for the voiced plosives.

2 English [p t k] have frequencies of 1.78, 6.42 and 3.09, while [b d g] have frequencies of 1.97, 5.14 and 1.05 (Gimson 1989: 219). In a set of two coordinates, with frequency of occurrence along the y-axis and the three places of articulation on the x-axis, draw two graphs, one for the voiceless plosives and one for the voiced plosives. Explain why the positions of [p] and [g] differ from those in the ‘theoretical’ graphs of the previous question.
3.4 CONCLUSION

Obviously, the fact that languages show so many similarities in their sound structures cannot be accidental. Different perspectives have been taken on this fact. The first explicit response has been an assumption that linguistic structure is an inherent, ‘innate’ property of human beings, whereby the shape and form of the innate component is yet to be determined. We are uncertain of the extent to which phonologists believe that certain segments or particular distinctive features are innate, but our guess is that most would maintain that it is contrastiveness which is innate, but that the way contrasts get spelled out in features may not be. The most general statement of the aims of phonology is that it seeks to establish the ‘possible space’ of phonological structure, and to show that the actual phonological systems we find in languages fit into that space, while non-existent structures do not. There are, for instance, many processes, but the number of possible processes that are never attested is very much larger. Adherents to the innateness position will argue that innateness may explain why children learn phonological structures so quickly. Humans may start out with a certain amount of ‘skeletal’ information, which they fill in with language-specific information on the basis of the language they are exposed to. With regard to syllable structure, for instance, the innate information might be that there are syllables, and that syllables have peaks and onsets. What the child would want to know next is (1) whether the onset may remain empty and/or (2) whether there may be a coda (Clements and Keyser 1983: 29; Kaye 1989: 56; Blevins 1995). At the other end of the innateness spectrum, the assumption will be that phonological structure is ‘emergent’. Given the input language and given a host of ambient factors, including the physiological properties of speakers and hearers that were discussed in chapter 2, phonological feature systems, segment inventories, syllable structures and so on arise and develop. It seems evident that both positions are correct. The issue is to understand just how we employ our innate cognitive and physiological structures to learn and use languages, and what the role might be of any more peripheral social and climatic factors.

A crucial assumption that underlies the notion of a phonological system is that the pronunciation of a language can be described with the help of a finite set of discrete constituents, i.e. segments, syllables, feet, etc. The pronunciation of every morpheme consists of a particular configuration of those constituents. These phonological constituents are meaningless, and distinct from the meaningful, morphosyntactic constituents of the language, such as morphemes, morphological words, syntactic phrases and sentences. In chapter 12, we will see how phonological structure continues above the level of the word.

NOTE

1 Based on the UCLA Phonological Segment Inventory Database (UPSID), which can be conveniently approached with the help of Henning Reetz’s search program at http://web.phonetik.uni-frankfurt.de/upsid_info.html.
4.1 INTRODUCTION

In chapters 1 and 3, we saw how different languages have different phonologies. One of the clearest illustrations of this fact is provided by the adaptation of loanwords to the phonology of the borrowing language. In this process, speakers will interpret the pronunciation of the words of the foreign language in terms of the phonological elements of their own. The way in which they do this can tell us a great deal about the phonology of the speaker’s native language. For example, the French pronunciation [fiˈlin] for English [ˈfiːlin] reveals that French does not distinguish tense and lax vowels, and uses [i] for both [i:] and [ɪ]. Second, it places an accent on the last syllable, regardless of where the stress was in the original word. In this chapter we will discuss the process of nativization, and illustrate it mainly on the basis of English loans in Hawaiian and one Indonesian loanword in Konjo. These languages have very different phonologies, the phonology of English being much more complex than that of Hawaiian, in particular. After showing how the pronunciation of foreign words is shaped by the phonological structure of the native language, it is pointed out that the phonological representation of native morphemes, too, may need to be adjusted. This need may arise when morphemes are combined. If a language with the syllable structure (C)V(C) only allows a coda consonant in word-final position, something will have to be done whenever a consonant-initial suffix like [ka] is attached to a consonant-final base like [taf], since *[tafka] would be ill-formed. In order to describe phonological adjustments, two approaches have been adopted: rules and constraints. The difference between these two approaches is briefly explained and illustrated.

4.2 HAWAIIAN

First, we give a brief outline of Hawaiian phonology, based on Elbert and Pukui (1979). The syllable structure is characterized by the formula in (1).

(1) (C)V(V)

That is, syllables do not have codas, vowels can be long or short, and the onset contains maximally one consonant. The Hawaiian phoneme inventory is extremely small. It is given in tabular form in (2). The rows in the C-system stand for manners of articulation, while the columns stand for places of articulation, ordered from
labial to glottal. For the vowels, the rows stand for degrees of tongue height, while
the columns have the order ‘front unrounded’, ‘back unrounded’ and ‘back rounded’.

(2) V C
   i u p k ʔ
   e o m n
   a w l

Consecutive vowels are diphthongs if the second is higher than the first (e.g. [au] or [oi]), and long vowels if they are identical (i.e. [aa] is [aː]). In terms of (1), these vowel sequences are VV and are monosyllabic. Other sequences of vowels are divided over two syllables. A [j] is inserted between [ie] and a lower vowel, and [w] is inserted between [uo] and a lower vowel. The latter consonant is indistinguishable from the unpredictable occurrence of the [w] listed in (2). A sequence like [ua] is therefore equivalent to [uwa], and [ia] is pronounced [ija]. Possible Hawaiian words are [iwa] ‘nine’, [niʔihau] (geographical name), [honolulu] (geographical name), [ala] ‘road’, [ʔala] ‘fragrant’, [puʔohi] ‘chatter’, [kaukau] ‘admonish’ and [hoʔolauleʔa] ‘celebration’. Impossible words are *[tuʔa], *[plai] and *[kehunanal].

Q25 Explain why the last three items cannot be Hawaiian words.

4.3 ADJUSTMENT PROCESSES

4.3.1 The process of nativization

A speaker of Hawaiian speaking English might well wish you [mele kelikimaka] on
25 December. This utterance is the result of the interpretation of [mɛi kɹɪmsɹ].
The situation exemplified here is representative of what happens when speakers of
one language decide to speak another language without adopting any of the phono-
logy of that other language. When faced with the task of pronouncing an expres-
sion in a foreign language while using only the phonology of their native language,
speakers need to (a) interpret each of the segments in the foreign word in terms
of the native segment system; and (b) make sure that no strings arise that break
the syllable structure constraints or any other phonotactic constraints of their language.
These two types of processes should be seen as different parses, according to Silver-
man (1992). The first parse takes place at the Perceptual Level: the acoustic input,
or the acoustic image that the speaker has of the foreign word, is interpreted as a
string of native segments. In the case of Merry Christmas, this process must have
resulted in the segments in the second column of (3). The segments in the first col-
umn represent the English interpretation of this expression.
Compared to the English segments, Hawaiian has phonetically quite similar segments available in most cases. The fact that there is no [r]-type consonant accounts for the interpretation of [ɪ] as [l], while [k], being the only lingual nonsonorant consonant, is the best interpretation of [s]. Now notice that this string of segments (those in the second column) is not well-formed. In particular, it cannot be analysed as sequences of (1), with or without the optional elements in that formula. The aim of the second parse, referred to as the Operative Level, is to make the string of segments perceived at the Perceptual Level conform to the phonotactic constraints of the language. Most importantly, the segments will have to be accommodated by giving them legitimate positions in syllable structure. Rather than throwing the [l] in Christmas out, therefore, a vowel is inserted between [k] and [l]. As a result, both [k] and [l] are now single consonants in their onsets, as required by (1). Similarly, vowels are inserted after the second and third [k]. (The new vowel would appear to be a copy of a nearby vowel, but we will ignore this aspect.)

Silverman’s division of the process of nativization into a Perceptual Level and an Operative Level is convenient and allows us to see the distinction between the phonological elements (vowels, consonants) and the phonological structure containing the segments. Respectively, they may be thought of as the things that need to be packaged and the legitimate ways of packaging them. However, we cannot really separate the two processes. It is not the case that the segments produced by the perceptual parse must be accommodated by the phonological structure at all costs, because it is not always the most likely segmental interpretations that survive. Rather, it would appear to be the phonetic output of the new phonological representation which is evaluated for its similarity to the phonetics of the original realization. Instead of supplying a vowel in order to make the segment string conform to the native syllable structure, the language might therefore adjust the string of segments by replacing a consonant with one that is not too different from it but that can be accommodated without the addition of a syllable. For instance, German does not allow voiced obstruents in the coda, which do occur in English. In (4), a

<table>
<thead>
<tr>
<th>Input</th>
<th>Perceptual Level</th>
<th>Operative Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>ɪ</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>𝑘</td>
<td>𝑘</td>
<td>𝑘</td>
</tr>
<tr>
<td>ɪ</td>
<td>i</td>
<td>i</td>
</tr>
<tr>
<td>𝑘</td>
<td>k</td>
<td>k</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>m</td>
</tr>
<tr>
<td>ø</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>s</td>
<td>k</td>
<td>a (cf. (1))</td>
</tr>
</tbody>
</table>

(3)
[b] in the English input for a German loanword is ultimately replaced with a [p], because the presence of [b] is ill-formed in that position. A solution that the German speaker does not resort to is to supply a vowel after the [b] so that it can be preserved as an onset consonant (*[paba]).

(4) 

<table>
<thead>
<tr>
<th>Input</th>
<th>Perceptual Level</th>
<th>Operative Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>b</td>
<td>(to preserve final position of plosive)</td>
<td></td>
</tr>
</tbody>
</table>

Another case of an alternative choice of consonant comes from Konjo. In this language, a word must minimally consist of a foot, which in Konjo is disyllabic and has the stress on the first syllable. It has the vowels [i e a o u], and its syllable structure is (C)V(C). Word-finally, the coda can only be [ʔ] or [ɲ]. It has a full set of nasals ([m n ɲ ɳ]) as well as a lateral (Friberg and Friberg 1991).

Q26 Why are the phonological representations [men] and [leʔem] not possible Konjo words?

Konjo has adopted the Indonesian word [ləm] ‘glue’ (itself a loan from Dutch [leim]) as [leʔen]. At first sight, this may be a little surprising. Note that [ləme] would be a possible word; in fact, the form [lame] is an actual word: ‘tuber’. The form [leʔen] conforms to Konjo phonology in that it is disyllabic and does not have [m] word-finally. Here, the Perceptual Level analysis must be [ləm], but as in the case of the German loan from English, the first decision at the Operative Level was to replace an impossible coda consonant with a possible one that is similar: [m] is replaced with [ɲ] (see (5)).

(5) 

<table>
<thead>
<tr>
<th>Input</th>
<th>Perceptual Level</th>
<th>Operative Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>l</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>e</td>
<td>e</td>
<td>e</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>ɲ (to preserve final position of the nasal)</td>
</tr>
</tbody>
</table>

Sometimes, a segment that is present in the original form is not translated in the loanword, which may happen if the segment is not particularly salient. An example is the Cantonese word [lɪp] ‘lift’, an adaptation of English [lif]. Here, while the first three segments in the English original have been translated, the final [t] was simply left uninterpreted.

A quick way of learning something about the phonology of a language is to look at loanwords borrowed from languages that you do know the phonology of. The adjustments that are made will indicate what structures are ungrammatical in the borrowing language.
In this section it has been shown that the pronunciation of foreign words is adjusted to the phonological structure of the native language. In the next section we will see that in the native phonology similar adjustment processes may occur.

Q27  Japanese has the following processes, which apply additively.

1. The coronal plosive [t] is affricated to [ts] before [i u].
2. The coronal obstruent [s] is prepalatal [ɕ] before [i].
3. In casual speech, the close vowels [i u] are devoiced to [i ŭ] between voiceless obstruents or after a voiceless obstruent at the word end.

Give the narrow transcriptions of /tukemono/ ‘pickled vegetables’, /sitá/ ‘tongue’ and /ótiba/ ‘fallen leaves’.

Q28  To salvage consonants that would otherwise be illegitimate codas, Japanese provides the vowel [u] to allow them to be onsets, as in [tɕiːˈzu] ‘cheese’ and [mɑʂuku] ‘mask’. However, after [t] the vowel [o] is used, as in [tόːsuto] ‘toast’. Referring back to Q27, can you explain this fact?

Q29  Here are some Japanese loans from Dutch (Vos 1963; de Graaf 1990).

<table>
<thead>
<tr>
<th>Loanword</th>
<th>Dutch origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>soːda</td>
<td>‘soda’</td>
</tr>
<tr>
<td>koːhi</td>
<td>‘kofi’</td>
</tr>
<tr>
<td>korera</td>
<td>‘xoːlara’</td>
</tr>
<tr>
<td>meranukoːri</td>
<td>meːlaŋxoː ‘li’</td>
</tr>
<tr>
<td>hipokonduːri</td>
<td>hipoːxɔn’dri</td>
</tr>
<tr>
<td>mangan</td>
<td>maŋ’yaːn</td>
</tr>
<tr>
<td>bombon</td>
<td>bɔm’bɔn</td>
</tr>
<tr>
<td>gomu</td>
<td>ɲm</td>
</tr>
<tr>
<td>kari</td>
<td>‘kaːli’</td>
</tr>
<tr>
<td>kiniːne</td>
<td>ki’ ninɻ</td>
</tr>
<tr>
<td>mesu</td>
<td>mes</td>
</tr>
<tr>
<td>karuːku</td>
<td>ɭaɻk</td>
</tr>
<tr>
<td>masuto</td>
<td>mast</td>
</tr>
<tr>
<td>buriki</td>
<td>blik</td>
</tr>
<tr>
<td>orugoːru</td>
<td>‘ɔryːl</td>
</tr>
<tr>
<td>sukopːu</td>
<td>sxɔp</td>
</tr>
</tbody>
</table>
1 List the consonants that appear word-finally in the Japanese words.
2 List the consonants that can appear in the coda of non-final syllables.
3 Why is a vowel added after the word for ‘rubber’ but not after the word for ‘manganese’?
4 On the basis of these data, what would you say are the voiceless fricatives of Japanese?
5 What evidence is there that Japanese does not allow CC-onsets? (In fact, Japanese allows CC-onsets, provided the second C is [j], as in the geographical name [kjo:to]).

Q30 Study the following Hawaiian loanwords from English carefully.

<table>
<thead>
<tr>
<th>English</th>
<th>Hawaiian</th>
<th>English</th>
<th>Hawaiian</th>
</tr>
</thead>
<tbody>
<tr>
<td>'ælbet</td>
<td>?alapaki</td>
<td>‘Albert’</td>
<td>wain</td>
</tr>
<tr>
<td>'tikt</td>
<td>kikiki</td>
<td>‘ticket’</td>
<td>rais</td>
</tr>
<tr>
<td>səʊp</td>
<td>kopa</td>
<td>‘soap’</td>
<td>bel</td>
</tr>
<tr>
<td>bɔ</td>
<td>pia</td>
<td>‘beer’</td>
<td>'fləo</td>
</tr>
<tr>
<td>kilt</td>
<td>kiliki</td>
<td>‘kilt’</td>
<td>‘zɔodiək</td>
</tr>
<tr>
<td>braʃ</td>
<td>palaki</td>
<td>‘brush’</td>
<td>‘θoazənd</td>
</tr>
<tr>
<td>'stɔ:ri</td>
<td>kole</td>
<td>‘story’</td>
<td>pa:m</td>
</tr>
<tr>
<td>sku:l</td>
<td>kola</td>
<td>‘school’</td>
<td>‘ɛləfənt</td>
</tr>
</tbody>
</table>

1 List the onset and coda clusters in the input forms that are broken up into different syllables in the Hawaiian output forms.
2 List the consonant clusters that are not so broken up, i.e. that are not fully interpreted.
3 List the English consonants for which Hawaiian [p] is used. What articulatory aspect do these consonants have in common?
4 List the English consonants for which Hawaiian [k] is used. What articulatory aspect do these consonants have in common?
5 Can you make a guess as to what the nativized Hawaiian form for English *false* might be?
6 If someone asked you why the Hawaiian speaker shouldn’t simply say [s] instead of [k] in words like *soap* and *Christmas*, what would your answer be? Would it be correct to answer that the fully native speaker of Hawaiian cannot hear the difference between [s] and [k]?
4.3.2 Adjustments in the native vocabulary

It may seem self-evident that the phonological shape of the morphemes of a language will conform to the structural constraints holding in that language. By and large, this is true: a language that disallows complex onsets will typically not have words with complex onsets. Yet it frequently happens that ill-formed structures potentially arise in native forms. For one thing, the phonological shape of morphemes that cannot by themselves be words, like affixes, need not conform to the constraints on syllable structure. An affix may consist of a single consonant, for instance, and as such not be a good syllable. Usually, well-formed syllables are only required at the level of the word. More generally, when suffixes are attached to bases to form complex words, or when words come together to form phrases, there is no guarantee that the phonological shape of the combination is well-formed, and adjustments are therefore frequently called for. For example, the underlying form of the English nominal plural suffix is [z]. This particular phonological representation is fine when the morpheme attaches to *eye to form *[əz], but it cannot be used in that same form in combination with either *nose or back. In the first case, a vowel is inserted between the stem and the suffix, to form *[nəʊz]z*, as the form *[nəʊzz]* is ill-formed; in the second case, *[bæks]* is formed, it being impossible to have two adjacent obstruents in the same syllable that differ in voicing (*[bækz]*). As a result, the suffix [z] has three different pronunciations, or morpheme alternants, [z], [s] and [tʃ].

4.4 TWO APPROACHES

There have been two approaches to the question of how phonological adjustments of (native) morphemes should be described, one based on rules that change representations and one based on constraints that require representations to have certain forms.

4.4.1 Rules

The traditional approach, associated with the monumental work by Noam Chomsky and Morris Halle, *The Sound Pattern of English (SPE)* (1968), uses rules that change the phonological representation of the morpheme in particular phonological contexts. These rules are ordered: each rule except the first applies to the output of the preceding rule. After all the appropriate changes have been made, the correct form surfaces. This is a derivational approach, in the sense that the surface form of the expression is derived in a series of structure-changing operations from the underlying form.

In the case of the English plural suffix, the derivational approach postulates two rules. One, **i-insertion** (6), inserts a vowel between two sibilants ([s z ʃ]). The other, **devoicing** (7), devoices an obstruent after a voiceless obstruent in the same syllable.

(6) **i-insertion** Insert [ɪ] between two adjacent sibilants in the same word.
(7) **devoicing** A voiced obstruent becomes voiceless after a voiceless obstruent.
The adjustments are given in the derivation (8). The first line represents the underlying forms, and the last line the surface forms. Intervening lines show the work of the phonological rules, in the order in which they apply. Thus, rule (6), 1-INSERTION, does not apply to [bæk-z] and [ar-z] because these forms have no instances of adjacent sibilants, but it does apply to [kis-z], which contains a sequence of [z] and [s] in a single word. Non-application is sometimes explicitly indicated by (n.a.), as has been done here.

\[(8)\]
\[
\begin{array}{l}
\text{Input representations} \\
\text{Rule (6)} \\
\text{Rule (7)} \\
\text{Output} \\
\hline
\text{bæk-z} & \text{kis-z} & \text{ar-z} \\
\text{(n.a.)} & \text{i} & \text{(n.a.)} \\
\text{s} & \text{(n.a.)} & \text{(n.a.)} \\
\text{baeks} & \text{kis-iz} & \text{aiz} \\
\end{array}
\]

An important advantage of applying rules in sequence is that they can express generalizations in simple ways. Notice, for example, that DEVOICING can be formulated in the simple way that it has been in (7), thanks to the work of 1-INSERTION. If we were to apply DEVOICING before 1-INSERTION, it would change the [z] after [kis] into [s], because it appears after a voiceless obstruent in the input. By first inserting the [i], the [z] is no longer adjacent to [s] and is skipped by DEVOICING. Of course, in rule-based descriptions not all rules are crucially ordered in this way.

Q31 The following are the underlying forms of a number of words in Tonkawa (Phelps 1975).

netale-o?  ‘he licks it’
we-netale-o?  ‘he licks them’
netale-n-o?  ‘he is licking it’
we-netale-n-o?  ‘he is licking them’
picena-o?  ‘he cuts it’
we-picena-o?  ‘he cuts them’
picena-n-o?  ‘he is cutting it’
we-picena-n-o?  ‘he is cutting them’

1 Identify the underlying forms of the morphemes for ‘lick’, ‘cut’, ’3SG-SUBJ’, ’3SG-OBJ’, ’3PL-OBJ’, and the ‘PROGRESSIVE’ (cf. English to be [Verb]-ing). (NB: One of these does not have a phonological form in Tonkawa.)

2 Two rules produce the surface forms.
   (i) CONTRACTION Delete the second vowel in a word-initial sequence CVCVCV.
   (ii) TRUNCATION Delete the second of two adjacent vowels.

   Give underlying forms, rule applications and surface forms of ‘he licks it’, ‘he is cutting them’ and ‘he cuts them’, using the two rules CONTRACTION and TRUNCATION.

3 Are the two rules ordered?
4.4.2 Constraints

The model of serially ordered rules may be compared to a production line in a factory. Each stage in that line is devoted to some specific operation which is performed on the unfinished product, with the last stage delivering the end product. Over the last two decades, a different approach has been taken to the phonological adjustments. In this constraint-based approach, demands are put on the surface form, and any form that does not comply with these constraints is rejected in favour of a form that does. The model is therefore more like a bureau for quality control than a factory. The most successful constraint-based theory is Optimality Theory, often abbreviated as OT (Prince and Smolensky 1993; Kager 1999). This theory holds that constraints are universal. There are two important features of the theory that explain why languages nevertheless have different phonologies. First, languages differ in the importance they attach to the various constraints. That is, the phonology of a language is given by the ranking of the set of universal constraints, known as that language’s constraint hierarchy. Second, constraints may be contradictory, and thus may be violated: if two constraints are contradictory, the one that is ranked higher will have priority.

A constraint-based approach to the English nominal plurals might postulate constraints (9) and (10). According to (9), sequences of sibilants are excluded within the same word. And according to the constraint in (10), we cannot have a sequence of obstruents that differ in that one is voiced and the other is voiceless. Thus, a form like *[nəuzz] violates the first constraint, and a form like *[bækz] the second.

(9) *SinSib: Sequences of sibilants are prohibited within the word.
(10) *aVoice–aVoice: Sequences of obstruents within the syllable must agree for voicing.

So how does Optimality Theory determine what the output form must be? For any given input form, there will initially be an unlimited set of output forms. This free generation of potential output forms is taken care of by a function called Gen (for ‘Generator’), which is subject only to very general constraints of well-formedness. Let us illustrate this with the plural for [kis]. For an input form [kis-z], some of the generated outputs will be [kisz], [kisz], [kis], [kizz], but also forms like [pets] or [tæpt]. There are two general forces at work that determine which of these numerous potential output forms is chosen by the language. One of these forces is called faithfulness: it is the force that tries to make the output form identical to the input form. Thus, if English were completely faithful, the plural of kiss would be [kisz]. The other force might be said to be the unmarked way of pronouncing things. If this force were allowed to have its way, unchecked by any other force, all words in the language, or indeed in all languages, would end up as something like [ba], or perhaps [ta]: anything more than this would be more ‘marked’ in the sense of less common, more complex and more difficult to pronounce or perceive. In reality, the outcome is determined by how these two forces interact. Each of the forces is represented by a set of universal constraints, and every language ranks these constraints in its own way. Again, if all the Faithfulness constraints are ranked
above all the phonological constraints, no phonological adjustments will be made to the input form. However, typically one or more phonological constraints are ranked above one or more Faithfulness constraints, which means that in the case of a conflict, the phonological constraint wins. Every constraint that is inspected will thus throw out a number of candidate forms, and this process goes on until there is only one form left. Optimality Theory thus holds that the output form is the **optimal form**, the form that is left as the only survivor of all candidate forms after an inspection of the constraint hierarchy.

McCarthy and Prince (1993) propose three important constraints to express Faithfulness. Max-IO requires that each segment in the input form (‘T’) has a corresponding segment in the output form (‘O’). That is, the input is ‘maximally’ represented in the output, and the constraint is therefore violated if a segment is deleted. Dep-IO requires that each segment in the output form has a corresponding segment in the input form. That is, the output must be entirely ‘dependent’ on the input, and the constraint is violated by any inserted segment. Third, Ident(F) requires that every feature (‘F’) of the input segment is ‘identical’ to every feature in the output segment. That is, this constraint is violated if a segment changes from voiceless [t] to voiced [d], say, or from bilabial [m] to dorsal [n]. (The theory allows these constraints to be split up into detailed subconstraints, and they are therefore better seen as constraint families.) There are many phonological constraints, some of which we will present informally below. It will be clear that the output form will be as close as possible to the input form, and that every deviation must be forced by some higher-ranking phonological constraint. It will thus be clear that an output [tæpt] for an input [kisz] is unlikely to survive an evaluation by the Faithfulness constraints, which will quickly see to it that [tæpt] is discarded in favour of forms that make a better job of preserving the input.

(11) Max-IO: Deletion of segments is prohibited.

(12) Dep-IO: Insertion of segments is prohibited.

(13) Ident(F): A segment in the input is identical to the corresponding segment in the output.

The operation of evaluating the collection of possible output forms is called Eval (for ‘Evaluation’). This evaluation is shown in **tableaux**. The tableau in (14) will serve as an illustration. The constraints are arranged in the columns, and the forms to be evaluated are arranged in the rows. The input form to be evaluated is given in the top left corner. A * in a cell indicates that the form of that row breaks the constraint in that column, and *! indicates that such a violation eliminates that form from further consideration: the violation is fatal. The optimal form, the winner, is marked $\square$. Shaded cells indicate that the constraint in that column has become irrelevant to the fate of the form in the row concerned. As is shown in (14), it is more important in English to obey *SibSib than to obey Dep-IO: in order to prevent the adjacency of [s] and [z], a segment [ɪ] is inserted between them. However, as shown in tableau (15), the language is less concerned about *ɑvoice–ɑvoice: although it could have saved the voiced [z] from ‘becoming’ [s] in [bækz] by inserting a vowel to form [bækɪz], it chooses not to break Dep-IO for this purpose. To
satisfy *a\text{voice–a\text{voice}}, it is, however, prepared to violate Ident(F) in the matter of the voicing of [z]: the input [z] corresponds to a non-identical output [s].

Incidentally, in addition to (9) and (10), there would have to be a constraint requiring that the phonological content of the stem must be retained. Otherwise, *[bægz], which satisfies *a\text{voice–a\text{voice}}, would also be described as a correct form, by the side of [bæks].

**Q32** The form *[kìs] is an incorrect output form for the plural *kisses.*

1. Which constraint (not listed in tableaux (14) and (15)) is responsible for ruling this form out?
2. Draw a tableau for the input form [kìsz] (i.e. the plural of *kìss*), with the two potential output forms [kìs] and [kìsz]. There should only be two constraints in your tableau: Dep-IO and the constraint you gave as the answer to the previous question. Would you rank the latter constraint above or below Dep-IO? Motivate your answer.

**Q33** The consonant inventory of Mauritian Creole contains the coronal fricatives [s \text{z}]. Palatoalveolar fricatives are absent. French words containing such fricatives, like [ʃəvø] ‘hair’ and [ʒenəræl] ‘general’, are adapted as [seve] and [jëve], respectively. Assume the two constraints Ident(F) and *ʃ/ʒ. Draw a constraint tableau with the correct ranking of these two constraints which shows the fate of the possible output forms [seve] and [jëve] in Mauritian Creole.

## 4.5 CHOOSING BETWEEN RULES AND CONSTRAINTS

You may now think that there is not much difference between the two approaches. There are, however, two important differences. First, in a derivational approach,
constraints and rules will sometimes both be needed and, moreover, appear to do the same work. Crucially, a rule-based grammar will also need constraints to characterize the phonological well-formedness of morphemes that have only a single form. For instance, the constraint *avoice–avoice (10) is generally needed to characterize the well-formed syllables of English: *[æbs], *[æpz], *[zti:] and *[sbi:] are ungrammatical, while *[ks] and *[zed] are fine, and are indeed used for the words ox and adze, as would be [bæks], the surname of composer Arnold Bax. However, because the plural form [bæks] needs to be derived from [bæk-z], a rule of devoicing (7) is required to effect the change from [z] to [s]. That is, the rule-based approach needs both (7) and (10), even though they would appear to be describing the same regularity of English. This inherent drawback of the rule-based approach is known as the duplication problem (Kenstowicz and Kisseberth 1977: 136). A constraint-based approach does not run into this problem: the constraint on English syllable structure will act as a condition on all forms, regardless of whether they are morphologically derived or not.

A second difference concerns the phonological adjustments that only appear in loanwords. For the purposes of the native phonology, Hawaiian does not require a rule that breaks up consonant clusters, since all the morphemes of the language conform to the requirement of the structure (C)V(V): there simply are no morphemes that begin with consonant clusters. In order to account for the adjustment made when a word containing a complex onset like [kr] is adapted to the native phonology, the derivational approach will have to add such a rule to the phonology (in this case, one to change [r] into [l], and one to insert a vowel between the [k] and the [l]). In a constraint-based description, no such ‘extra’ grammar would need to be supplied. To characterize the native Hawaiian forms, a constraint forbidding complex onsets would need to be undominated anyway, meaning that no constraint ranks above it, since the language allows no exceptions. That is, in general, in an OT description, the constraints would be ranked so as to bring any new input forms in line with the structural demands of the language.

To illustrate this last point, consider again the Konjo word for ‘glue’, [leʔen]. The input form is the Indonesian word [lem]. Among the generated output forms there will be forms like [lem], [lemem], [leʔ], etc. Why isn’t the output [lem]? This is because the language has phonological constraints that are ranked higher than the Faithfulness constraints. One such constraint is that the output should minimally be a disyllabic foot, a constraint that is undominated in the language, so that there are no exceptions.

(16) MinWord: A word is minimally a (binary) foot.

Since MinWord outranks Dep-IO, a disyllabic form will be preferred to the monosyllabic [lem]. Let us suppose that we have evaluated the potential outputs up to a point where all monosyllabic candidates (e.g. [len], [lem], [leʔ]) have been discarded by MinWord. There are two further constraints that are relevant at this point. One, CodaCondition (17), forbids the presence of consonants other than [ŋ ʔ] in the coda.
(17) CODACond: A coda consonant is [ʔ] or [ŋ].

After inspection of CODACond, there are two reasonably faithful candidates left: [leʔeŋ] and [leme]. (We will ignore the question why [ʔ] and [e] are the best choices here for the inserted segments.) The second constraint that is relevant at this point is ALIGN-STEM-RIGHT (18), which requires that no segments should be added to, or removed from, the end of the input. Formally, the requirement is that the end of [lem] should coincide with the end of a syllable in the output. Alignment constraints are highly serviceable in OT and are generally responsible for locating both morphological and phonological constituents in the linguistic structure. For instance, a prefix is described as a morpheme that aligns its left edge with the left edge of the word it is included in (McCarthy and Prince 1993).

(18) ALIGN-STEM-RIGHT: The last segment of the input corresponds with the last segment of the output.

In Konjo, ALIGN-STEM-RIGHT outranks DEP-IO. The form [leʔeŋ] passes the higher-ranked constraint, because the end of the last syllable, [ʔeŋ], coincides with the end of the input form [lem], and [leme] fails it, because the end of the syllable [me] does not. These correspondences are shown in (19). (We could of course, perversely, assume different correspondences. For instance, the output [m] in (19a) might be taken to correspond to the second [e] in the input, but that output form would be thrown out on the grounds of multiple violations of specific versions of IDENT(F). Not even the consonantal nature of [m] would have been preserved, and an assumption of a high-ranking IDENT([consonant]) would be enough to discard that output form.)

(19) a. l e ʔ e ŋ b. l e m e
    l e m l e m

The fact that [leʔeŋ] fails the lower-ranking DEP-IO is of no importance: since there are no other candidates left, it is irrelevant whether the successful candidate breaks any constraints ranked below the constraint that eliminated the last rival form(s). Neither is it of any relevance that forms that failed to pass a constraint satisfy any lower-ranked constraints. The tableau in (20) schematizes the situation.

(20)  

<table>
<thead>
<tr>
<th></th>
<th>ALIGN-STEM-RIGHT</th>
<th>DEP-IO</th>
</tr>
</thead>
<tbody>
<tr>
<td>lem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leme</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>≠ leʔeŋ</td>
<td>*</td>
<td>**</td>
</tr>
</tbody>
</table>

In other words, the adjustment of Indonesian [lem] to Konjo [leʔeŋ] is accounted for with the help of the same constraint hierarchy as is used for characterizing the native forms of the language (Yip 1993; Jacobs and Gussenhoven 2000). It would therefore appear that, at least in principle, OT scores over a description that uses structure-changing rules in two ways. First, OT can characterize the phonological
grammatical by making the form fit in a way that avoids having to state the same information twice, once as a constraint on monomorphemic forms and once in terms of a rule that changes polymorphemic forms so as to conform to that constraint. Second, to account for phonological adjustments that are only observed in the adaptation of loanwords (i.e., never in the native forms), a rule-based analysis must add rules to the native grammar that bring those adaptations about. By contrast, the constraints that in an OT analysis characterize the native forms should equally characterize the optimal form of any incoming loanword.

Q34 In Japanese words, only a single voiced obstruent may occur. This constraint is known as Lyman’s Law. As a result, [pato], [bato] and [pado] are possible words, but *[bado] is not.

1 Explain why *[gazi] isn’t, but [gami] is, a possible word.

Compounds are formed from two words, like English armchair from arm and chair. Each of the words making up a compound forms a separate domain for Lyman’s Law. In many compounds, the initial consonant of the second word is voiced, as shown in (1). This is known as Rendaku (Ito and Mester 1986).

(1) ori ‘coloured’ kami ‘paper’ ori-gami ‘coloured paper’
    take ‘bamboo’ sao ‘pole’ take-zao ‘bamboo pole’
    garasu ‘glass’ tama ‘beads’ garasu-dama ‘glass beads’

Rendaku fails systematically in the compounds illustrated in (2).

(2) kita ‘north’ kaze ‘wind’ kita-kaze ‘north wind’
    tsuno ‘horn’ tokage ‘lizard’ tsuno-tokage ‘horned lizard’
    siro ‘white’ tabi ‘tabi’ siro-tabi ‘white tabi’

2 What might be the reason that Rendaku fails in (2)?

3 If you were to give an OT analysis of compounds, and Rendaku and Lyman’s Law are two constraints, how would you rank them? Put the two constraints in a tableau with /kita-kaze/ as input, and *[kita-gaze] and [kita-kaze] as outputs.

4.5.1 Gradient violation and unranked constraints

Sometimes, two forms break the same constraint, but one of them can be said to violate the constraint more than the other form. Also, it will frequently happen that two constraints are never contradictory and therefore that their ranking makes no
difference. We illustrate the first point with the help of a hypothetical language, Un-Konjo, in which the ranking of ALIGN-STEM-RIGHT and DEP-IO is the reverse of that in Konjo. This has been done in tableau (21). Notice that inspection of the form [leʔen] leads to two violations of DEP-IO, compared with only one in the case of [leме]. Multiple (or ‘gradient’) violation is shown by the number of stars in the relevant cells. Since the form [leʔen], with two inserted segments, is a worse candidate for the purposes of this constraint than [leме], which has only one inserted segment, it is eliminated. Thus, [leме] is the winner in Un-Konjo, and ALIGN-STEM-RIGHT, along with all other lower-ranked constraints, is left uninspected.

Let us go back to Konjo. Observe, again, that the potential output form *leʔem is non-optimal because it breaks CODACOND. Like ALIGN-STEM-RIGHT, this constraint must therefore be ranked above IDENT(F), which requires that [m] remains [m]. However, the ranking of CODACOND with respect to the two constraints in tableau (20) is indifferent. We could insert it either before or after them. In tableau (22) we have placed it in first position; to indicate that its ranking is indifferent, it is separated from ALIGN-STEM-RIGHT by a dotted line. Likewise, while IDENT(F) is ranked below CODACOND and ALIGN-STEM-RIGHT, it does not interact with DEP-IO.

Q35 Draw two tableaux that are identical to tableau (22), except that CODACOND is, respectively, in third and in fourth position. Please motivate your answers to the questions below.

1. Is DEP-IO crucially ranked above CODACOND?
2. Is CODACOND crucially ranked above IDENT(F)?

These brief illustrations of the rule-based approach (section 4.4.1) and of a constraint-based approach (section 4.4.2 and this section) give only the main ideas behind the derivational theory of SPE and of Optimality Theory. In later chapters we will illustrate both approaches more extensively. In chapters 7, 8 and 12 we present various descriptions in the derivational framework of SPE, and in chapter 11 we give an account of the word-stress locations in a number of languages in the framework of Optimality Theory.
4.6 CONCLUSION

In this chapter we have seen that the phonologies of languages actively impose phonological adjustments on input forms. Such adjustments are most readily observable in loanwords. However, the need to make adjustments also arises in the native vocabulary of the language when morphemes are combined in words and phrases, since such combination may lead to the creation of phonological representations that are ill-formed. There are two ways in which phonological adjustments have been described. First, they can be described with the help of a series of rules, which successively change the representation so as to make it conform to the requirements of the language. Second, they can be described by means of output constraints that state what forms must look like. Optimality Theory is a constraint-based theory that postulates that a language can be characterized by a ranking of a set of violable, universal output constraints, and that the correct form is the form that violates the constraints least.
5.1 INTRODUCTION

In chapters 2 and 3 we saw that the segment inventories of languages can be divided into subgroups. Thus, we have separated the group of vowels from the group of consonants, and when we discussed the way languages ‘build up’ their inventories, we distinguished groups of voiceless obstruents (e.g. \[p t k\]) from voiced groups that are otherwise the same (i.e. \[b d g\] in our example). This would appear to suggest that the segment is not the smallest constituent of phonological structure. In this chapter, we will motivate this assumption. We will introduce the distinctive features that will in principle enable us to describe the segments in the world’s languages, and to refer to those groups of segments that play a role in their characteristic phonological processes and constraints. The latter consideration will be shown to provide an important motivation for the assumption of distinctive features. In this perspective, these features are the elements by which we can refer to natural segment classes, groups of segments that are treated as groups by languages.

The distinctive features discussed in this chapter are binary, meaning that they always have either a positive or a negative specification, like [+voice], as for \[b\], or [−voice], as for \[p\]. We introduce three groups of these features. The ‘major-class’ features compartmentalize segments into four classes by means of two binary features. Next, we discuss three binary features specifying states of the glottis (‘laryngeal’ features). Finally, there are four features with which we can specify manners of articulation (‘manner’ features). We will see that not all segments are specified for all features. That is, in addition to segments that are [+nasal] and [−nasal], there are segments that have no specification for nasality. The segment [ʔ], for instance, which is produced with the help of a glottal closure, is not specified for any manner feature.

5.2 MOTIVATING DISTINCTIVE FEATURES

Inventories of segments can be seen as composed of intersecting sets of segments, like ‘all the voiceless segments’ or ‘all the segments articulated with lip rounding’ or ‘all the segments with a coronal place of articulation’. This fact does not in itself imply that a segment is represented in the synchronic structure in terms of a number of separate features, rather than as an unanalysable constituent. It could be the case that we are simply dealing with a reflex of the way segment inventories developed historically. Many cathedrals in Europe contain
elements of earlier buildings, but this does not mean that these elements are in any sense functional today.

The chief motivation for the introduction, and hence definition, of a feature is that it enables us to characterize a natural segment class. It appears to be the case that languages frequently refer to particular groups of segments, while other conceivable groupings are never referred to. Thus, languages frequently ban voiced obstruents from the final position in the word or from the syllable coda, but no language would ban [m d ʒ] from the coda while allowing [b n ɣ]. If particular groups of segments figure again and again in phonological generalizations about syllable structure or contextual variation, then evidently those groups must share something. This is what we intend to capture with the help of features by which phonological grammars can recognize them.

Q36
1 Which consonants are aspirated syllable-initially in English?
2 If we ignore the ordinal suffix [θ], as occurring in sixth, which consonants can form a word-final complex coda in a position after [s] in English?
3 Which consonants can occur between [s] and [r] in the syllable onset of English words?
4 What is the significance of the fact that the preceding three questions have the same answer?

In addition to the requirement that distinctive features should enable us to refer to natural segment classes, we should require of a feature analysis that the distinctive features can characterize the segment inventories of the languages of the world. That is, all segments must be characterizable in terms of some unique combination of features. This requirement is quite self-evident: we don't want to end up with a list of features that cannot characterize the difference between [m] and [n], say.

There is a third requirement placed on distinctive features. Consider first that there is no a priori reason to suppose that natural segment classes should consist of phonetically similar segments. Instead of [p t k], it might have been the case that it was [p tʃ] which were aspirated in English. If more and more natural segment classes were found to be phonetically arbitrary, we might be led to believe that a distinctive feature was an abstraction, i.e. that it couldn't be given a definition other than in terms of the collection of segments that had it. To continue the hypothetical example, we would say that [p tʃ] are [+delta], and that [+delta] consonants are aspirated in some context. In reality, we see time and again that the segments in natural segment classes are phonetically similar. This has led to the requirement of the Naturalness Condition (Postal 1968: 73), according to which distinctive features must have a phonetic (articulatory or acoustic) definition. Notice that the relationship is not the other way around. We need not, indeed should not, postulate
a distinctive feature purely because there is some phonetic property that a group of segments has in common. As Kaye (1989: 27n.) puts it: ‘One could group sounds according to the total energy involved in their production, the number of different muscles involved in their articulation, their length in milliseconds, the distance an articulator moves from some predefined neutral position, and so on.’ Obviously, none of those definitions corresponds to a natural segment class. The three requirements we must impose on a distinctive feature system, therefore, are that:

1. They should be capable of characterizing natural segment classes.
2. They should be capable of describing all segmental contrasts in the world’s languages.
3. They should be definable in phonetic terms.

5.3 FEATURE VALUES

Features may have values. A binary feature has either the value ‘+’ or the value ‘−’. The claim here is that both the group of segments that has the minus value and the group that has the plus value form natural classes. For instance, the assumption of the binary feature [±voice] implies that languages refer to groups of voiceless segments as well as to groups of voiced segments. Ever since they were first proposed in Jakobson et al. (1952), distinctive features have standardly been assumed to be binary. In recent years, phonologists have proposed univalent features, effectively a feature without a value (e.g. Ewen 1995). In this case, reference can only be made to the class of segments that has the feature, not to the collection of segments that does not possess it. Other terms for univalent are unary or privative. In chapter 6 we will discuss a set of univalent features characterizing the four main articulation zones, like labial. Multivalued features are no longer used in phonology. One that was common in the heyday of SPE was the feature [nstress], where \( n \) was a number to indicate the degree of stress of the vowel that was specified with the feature. Setting up and motivating a feature set is a core research task in the phonology of segments. Our feature set largely follows Halle and Clements (1983), but it has also been informed by Sagey (1986). Below, we discuss major-class features, which classify segments into segment types like ‘vowel’ and ‘obstruent’; laryngeal features, which specify the glottal properties of the segment; and manner features, which specify the type of constriction, or more generally the manner of articulation.

5.4 MAJOR-CLASS FEATURES

There are three major-class features, [±consonantal], [±sonorant] and [±approximant].

1. [±consonantal]. [+cons] segments have a constriction somewhere along the centre line in the vocal tract which is at least as narrow as that required for a
Distinctive features

67

fricative; [−cons] segments lack such a constriction. Thus, [+cons] are plosives, affricates, fricatives, nasals, laterals and [r], while [−cons] are vowels, glides like [jɥw], and, because their stricture is in the larynx rather than the vocal tract, [h ɦ ℓ]?

2 [±sonorant]. This feature distinguishes obstruents ([−son]) from sonorant consonants and vowels ([+son]). [+son] segments are produced with a constriction in the vocal tract which allows the air pressure behind it and in front of it to be relatively equal, while this is not the case for [−son] segments. That is, either [−son] segments have an oral constriction which causes a significant increase in the air pressure behind it (e.g. [s], [d]), or there is no constriction in the vocal tract. Since the vocal tract does not include the larynx, [h] and [ʔ] are [−son]. So [+son] are all vowels, glides like [v w j], liquids and nasals, while [−son] are plosives, fricatives, affricates and laryngeal segments.

Q37

1 How many (possibly overlapping) natural segment classes can be referred to with two binary features at one’s disposal? Hint: A natural class can be captured by one feature or by a combination of features.

2 Use [±consonantal] and [±sonorant] to characterize four classes, giving examples of segments for each class.

Q38 In Dutch, there is a rule that places a [ə] between the noun stem and the diminutive ending [tjə], as when [bal] ‘ball’ is affixed with [tjə] and becomes [’balətjə]. On the basis of the following data, characterize the group of segments after which this [ə] is inserted.

<table>
<thead>
<tr>
<th>bal</th>
<th>balətjə</th>
<th>‘ball’</th>
<th>dɪŋki təj</th>
<th>dɪŋki təjtjə</th>
<th>‘Dinky Toy’</th>
</tr>
</thead>
<tbody>
<tr>
<td>kom</td>
<td>kəmətjə</td>
<td>‘bowl’</td>
<td>kəp</td>
<td>kəpjə</td>
<td>‘hood’</td>
</tr>
<tr>
<td>laχ</td>
<td>laχə</td>
<td>‘laugh’</td>
<td>bɛs</td>
<td>bɛʃə</td>
<td>‘berry’</td>
</tr>
<tr>
<td>kan</td>
<td>kanətjə</td>
<td>‘jag’</td>
<td>rək</td>
<td>rəkə</td>
<td>‘skirt’</td>
</tr>
<tr>
<td>kar</td>
<td>karətjə</td>
<td>‘cart’</td>
<td>dɪŋ</td>
<td>dɪŋətjə</td>
<td>‘thing’</td>
</tr>
<tr>
<td>pet</td>
<td>petə</td>
<td>‘cap’</td>
<td>sək</td>
<td>səkə</td>
<td>‘sock’</td>
</tr>
</tbody>
</table>

Q39 In the variety of Spanish spoken in the state of Cordoba in Colombia, the first of two adjacent consonants was assimilated to the second, creating a geminate consonant (Charette 1989). For example, the word for ‘door’, which is [pwerta] in Peninsular Spanish, is [pwetta]
in the Cordoba variety. The process did not always apply. Characterize the class of consonants that underwent the process.

<table>
<thead>
<tr>
<th>Earlier form</th>
<th>Later form</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>serdo</td>
<td>seddo</td>
<td>‘pork’</td>
</tr>
<tr>
<td>awto</td>
<td>awto</td>
<td>‘car’</td>
</tr>
<tr>
<td>talko</td>
<td>takko</td>
<td>‘talc’</td>
</tr>
<tr>
<td>doktor</td>
<td>dottor</td>
<td>‘doctor’</td>
</tr>
<tr>
<td>algo</td>
<td>aggo</td>
<td>‘something’</td>
</tr>
<tr>
<td>neptuno</td>
<td>nettuno</td>
<td>‘Neptune’</td>
</tr>
<tr>
<td>fohforo</td>
<td>fohforo</td>
<td>‘match’</td>
</tr>
<tr>
<td>magdalena</td>
<td>maddalena</td>
<td>‘Madeleine’</td>
</tr>
<tr>
<td>ojgo</td>
<td>ojgo</td>
<td>‘onion’</td>
</tr>
<tr>
<td>arma</td>
<td>amma</td>
<td>‘weapon’</td>
</tr>
<tr>
<td>ahno</td>
<td>ahno</td>
<td>‘donkey’</td>
</tr>
</tbody>
</table>

**Q40** In Dutch, sequences of identical segments, which arise when the last segment of one morpheme is the same as the first of the next, are degeminated. Characterize the class of sounds that is subject to the rule in terms of distinctive features.

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>χe:l lampjə</td>
<td>χe: lampjə</td>
</tr>
<tr>
<td>fis sapjə</td>
<td>fis sapjə</td>
</tr>
<tr>
<td>le: χa:tjə</td>
<td>le: χa:tjə</td>
</tr>
<tr>
<td>fra:j jaχt</td>
<td>fra:j jaχt</td>
</tr>
<tr>
<td>lo:k kint</td>
<td>lo: kint</td>
</tr>
<tr>
<td>slim meiʃjə</td>
<td>slim meiʃjə</td>
</tr>
<tr>
<td>ryv ve:r</td>
<td>ryv ve:r</td>
</tr>
<tr>
<td>do:f fentjə</td>
<td>do: fentjə</td>
</tr>
<tr>
<td>χutko:p pak</td>
<td>χutko: pak</td>
</tr>
<tr>
<td>fei nɔ:jjə</td>
<td>fei nɔ:jjə</td>
</tr>
<tr>
<td>niu vəntjə</td>
<td>niu vəntjə</td>
</tr>
<tr>
<td>fi:r raːmən</td>
<td>fi: raːmən</td>
</tr>
</tbody>
</table>

3  [$±$approximant]. [+approx] are those segments which have a constriction in the vocal tract which allows a free (frictionless) escape of air, while for [−approx] segments this is not the case (Ladefoged 1971: 46; Clements 1993). Vowels and non-nasal sonorants, like [ɹʌ], are [+approx] segments. (The term ‘lateral’ is used for any l-type sounds, while the term ‘rhotic’ refers to any r-type sound; laterals and rhotics are often referred to as ‘liquids.’)
In many languages, the difference between [i u] and [j w] can be interpreted in terms of syllable constituency rather than phonological content. In such cases, [u] is different from a labial-velar approximant [w] not because these segments have different features, but because they occupy different positions in the syllable: the peak for [u], and the margin, usually the onset, for [w]. See also Q83.

### 5.5 LARYNGEAL FEATURES

There are three laryngeal features, [±voice], [±spread glottis] and [±constricted glottis].

1. [±voice]. [+voice] are segments for which the vocal folds are close enough together to allow vibration, while for [−voice] this is not the case. Thus, [+voice] are vowels (e.g. [i ë ã]), sonorant consonants (e.g. [m n l r w]) and voiced obstruents (e.g. [b z ã ð] and [f ã]), while [−voice] are voiceless obstruents (e.g. [p ð ñ õ t s h]).

2. [±spread glottis]. [+spread] segments have a vocal fold configuration that produces audible friction in the glottis, while [−spread] segments lack
such a configuration. Thus, aspirated segments like [pʰ kʰ] and [h fi] are [+spread], while other segments are [−spread].

3 [±constricted glottis]. For [+constr] segments the vocal folds are tense and drawn together, while for [−constr] segments this is not the case. Thus, [ʔ], laryngealized vowels (e.g. [u]) and laryngealized sonorant consonants (e.g. [mʔ]), glottalized obstruents (e.g. preglottalized [ʔp] or ejective [p’]) are [+constr]. So are implosives ([ɓ ɗ ɠ]). Other segments are [−constr].

Q43

1 American English [p t k] are accompanied by a glottal closure when appearing in the syllable coda, as in *sit, atlas, popcorn, duckpond*. What feature specification do these plosives acquire in this context?

2 In Southern Oromo, a rule of i-epenthesis inserts [i] between the ejectives [t’ tʃ] and a following [t n], as shown in (a). However, [ʔ d’t] do not trigger the rule but undergo other changes that are not relevant here, as shown in (b,c,d) (Lloret 1995). What combination of features distinguishes the [t’] from the three consonants that do not trigger i-epenthesis?

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>fit’-na</td>
<td>fit’ina</td>
</tr>
<tr>
<td>b</td>
<td>aʔ-na</td>
<td>a:na</td>
</tr>
<tr>
<td>c</td>
<td>fe:ɗ-ta</td>
<td>fe:ɗ:ta</td>
</tr>
<tr>
<td>d</td>
<td>bit-ta</td>
<td>bit:a</td>
</tr>
</tbody>
</table>

5.6 MANNER FEATURES

There are four manner features, [±continuant], [±nasal], [±strident] and [±lateral].

1 [±continuant]. [+cont] segments lack a central occlusion in the vocal tract, while [−cont] segments are produced with such an occlusion. Thus, plosives (e.g. [p d g]), nasal consonants (e.g. [m n]), affricates (e.g. [tʃ]) and laterals (e.g. [l]) are [−cont], and other segments are [+cont]. Some languages apparently treat laterals as [+cont], which is phonetically understandable in the sense that while these segments have a central occlusion, they have a lateral aperture.

2 [±nasal]. [+nas] segments (e.g. [m n]) are produced with the velum (‘soft palate’) lowered, and [−nas] segments have the velum in its closed (raised) position. Nasal consonants and nasalized vowels are [+nas]; other segments are [−nas].

3 [±strident]. [±strident] is relevant for obstruents only and refers to a type of friction. [+strident] segments cause a noisier kind of friction than [−strident]
Distinctive features

71

segments. [+strident] voiceless fricatives are [f s ʃ χ]; [−strident] ones are [θ ç]. Together with [CORONAL] (section 6.2.2), the feature can be used to capture sibilants ([s z ʃ tʃ dʒ]), needed for a correct description of the context for English i-INSERTION. Languages for which a contrast between [f v] and [φ β] has been reported include Ewe and Venda (Ladefoged and Maddieson 1996: 140). The name Ewe, [ɛβɛ], forms a minimal pair with the word for ‘two’, [ɛvɛ], in that language, while English contrasts [s z] with [θ ð], as in sigh, xi (the Greek letter ξ, [zaɪ]), thigh, thy. The feature’s other task is to distinguish plosives from affricates, both of which are [−son, −cont]. Such contrasts are common, as in German [tɔl] ‘mad’ vs [tsɔl] ‘import duty’, English [tæm] time vs [tʃæm] chime, or Corsican [ˈɟalu] ‘freeze+1sg’ vs [ˈdʒalu] ‘yellow’. The representation of affricates is controversial. A widely supported view, however, is that they are [+strident] plosives (Rubach 1994; Clements 1999). For one thing, it is at least suggestive that affricates typically have strident friction after the release of the closure, as in [pf ts tj kʃ] rather than [pφ tθ cç kx].

Q44 In Scottish English, [i e a o u ʌ ai] are pronounced as long [iː eː aː oː uː ʌː ai] in open syllables. The long vowels (including [ae]) also appear before certain consonants. How can this class of consonants be characterized? This regularity, described by Aitken (1981, 1984), is known as Aitken’s Law.

| rað | wriðe | mæl | mɪl |
| læin | nɪn | bɛːz | beɪg |
| tiːz | tɛs | rɔd | rɔd |
| læːv | lɔv | kɑːr | ɔr |
| lɪθ | Leɪθ | hɔm | hɔm |
| lɛf | lɪf | pɪs | pɪs |
| mel | mɪl | ræʃ | ræʃ |

Q45 In Turkish, obstruents are voiceless in the syllable coda, as shown in (1) (after Kim 1997).

<table>
<thead>
<tr>
<th>(1)</th>
<th>Underlying</th>
<th>Objective</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>ip</td>
<td>ipi</td>
<td>ipler</td>
<td>‘rope’</td>
</tr>
<tr>
<td>dib</td>
<td>dibi</td>
<td>dipler</td>
<td>‘bottom’</td>
</tr>
<tr>
<td>at</td>
<td>atu</td>
<td>atlar</td>
<td>‘horse’</td>
</tr>
<tr>
<td>ad</td>
<td>adu</td>
<td>atlar</td>
<td>‘name’</td>
</tr>
<tr>
<td>kɔk</td>
<td>kɔki</td>
<td>kɔkler</td>
<td>‘root’</td>
</tr>
<tr>
<td>gɔg</td>
<td>gɔgi</td>
<td>gɔkler</td>
<td>‘sky’</td>
</tr>
</tbody>
</table>
Distinctive features

72

4 [±lateral]. [+lat] segments have a central tongue contact in the oral cavity with one or both sides of the tongue being held away from the roof of the mouth, allowing the air to escape there, like alveolar [l] and prepalatal [ʎ]. Other sounds are [−lat]. A lateral escape of the air is also possible for obstruents, like the lateral fricatives [l] (voiceless) and [ʎ] (voiced) and the lateral affricates [tl] and [dʎ].

5.7 AMBIGUITY AND NON-SPECIFICATION

In this section we further illustrate the kind of reasoning by which featural analyses of segments are arrived at. As you will recall from section 5.2, the first and foremost motivation for the definition of a feature is that it allows natural classes to be referred to, as has been illustrated in the exercises so far in this chapter. It is possible, of course, that there is conflicting evidence for the inclusion or exclusion of a particular segment in a natural class. This is the case for [l], which behaves ambiguously vis-à-vis the feature [±continuant] when different languages are considered.
We will conclude that languages may differ in the feature value of [continuant] for [l]. A second topic discussed here is featural underspecification. It is to be expected that certain (classes of) segments are not specified for all features. That is, features may be irrelevant for certain segments. Thus, vowels will not be specified for [±lateral], simply because those segments cannot be differentiated with the help of this feature. Less obviously, it appears that [h] does not participate in rules referring to [±continuant]. In the section that follows, we will argue that laryngeal segments are not specified for either manner or place features.

5.7.1 The ambiguous behaviour of [l]

If you did Q44 correctly, you will have drawn the conclusion that in Scottish English [l] is [−cont]. In other languages, however, the same segment may have to be analysed as [+cont]. In Frisian, vowels are nasalized before [n] in the same syllable, provided a [+cont] consonant follows. The [n] itself is subsequently lost (Tiersma 1985). From the data in (1), which consist of infinitival verb forms prefixed with [in-] or [oən-], it is clear that the [l] must be [+cont]: the group of consonants that does not allow the change to go through is [p t k g n], while the consonants that do allow it are [s f j v r l].

\[
\begin{array}{lll}
\text{(1)} & \text{in-} & \text{impakə} \\
& \text{oən-} & \text{oəntrekaə} \\
& \text{oən-komə} & \text{oəŋkoma} \\
& \text{in-} & \text{ingən} \\
& \text{oən-nimə} & \text{oənnimə} \\
& \text{oən-stiən} & \text{oənnɪmə} \\
& \text{in-} & \text{fələ} \\
& \text{in-} & \text{iən} \\
& \text{in-} & \text{vəniə} \\
& \text{oən-} & \text{oəntrεkə} \\
& \text{in-} & \text{iλiə} \\
\end{array}
\]

\begin{align*}
\text{‘to wrap up’} \\
\text{‘to take to heart’} \\
\text{‘to arrive’} \\
\text{‘to enter’} \\
\text{‘to accept’} \\
\text{‘to please’} \\
\text{‘to fall in’} \\
\text{‘to give in’} \\
\text{‘to live with one’s parents’} \\
\text{‘to call’} \\
\text{‘to preserve’}
\end{align*}

5.7.2 Laryngeals have only ‘major-class’ feature specifications

The Frisian rule of nasalization may also give us an indication about another question that concerns the feature [continuant]. The laryngeal consonants are [−cons, −son]. If we take the definitions of these features seriously, this means that [h?] have a constriction in the larynx, but no constriction in the vocal tract, and thus do not have the sort of constriction that sonorants like [j] have. That is, they do not have a constriction in the vocal tract that is at least as narrow as used by fricatives. The question therefore arises whether it makes sense to want to specify [h?] for manner features or for place features, since, after all, if there is no stricture, how could we specify either its manner or place? While this makes good phonetic sense, the first question is whether languages make reference to any manner or place features of glottal consonants. For example, we might find rules that refer to sets of the type ‘[h] and other (classes of) [+cont] segments’. One such set is ‘all
vowels, the approximants [l r j w] and all fricatives, including [h], which would be the class [+cont]; another is [j w], vowels and [h] (but not [ʔ]), which would be the class [−cons, +cont]. As we have just seen, Frisian nasalization is a rule referring to the class of continuant consonants, but, interestingly, [h] is not among them, as shown in (2). If [h] were specified for manner features, that is, if [h] were [+cont], we would expect the rule of nasalization to apply before it. Since nasalization does not apply before [h], we have to conclude that this segment is not specified for manner features and hence cannot be considered [+cont].

(2) inhεljə ‘to hold in’
oənhιə ‘to listen to’

A Dutch assimilation rule points to the same conclusion. progressive devoicing devoices fricatives after obstruents, as shown in (3).

(3) ze: ‘sea’
    zər’naːl ‘journal’
    vyːr ‘fire’
    yas ‘gas’

    əp ‘se: ‘at sea’
    dt fər’naːl ‘this journal’
    kampfỳ:r ‘campfire’
    dt ‘xas ‘this gas’

The class of voiced fricatives could be referred to by [−son, +cont]. Clearly, if [h] is [+cont], this consonant would be included in that natural class. As it happens, the Dutch glottal fricative is frequently voiced. This voiced [ɦ] shows up after obstruents, too: stadhuis ‘town hall’ may be pronounced [stə’t’hojʌs] as well as [stə’t’hojɛs] (Rietveld and Loman 1985). Dutch [ɦ] is thus not subject to progressive devoicing and cannot be [+cont]. We now have evidence from two languages that [h] is not specified for continuancy. With McCarthy (1988) we will assume therefore that laryngeal segments are not specified either for manner or for place features.

Q47 In the Limburgish dialect of Geleen, the diminutive suffix [kə] is palatalized to [cə] if the stem ends in one of a group of coronal consonants.

<table>
<thead>
<tr>
<th>Stem</th>
<th>Palatalized Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>feːs</td>
<td>feːskə</td>
<td>‘feast, party’</td>
</tr>
<tr>
<td>mɪŋʃ</td>
<td>mɪŋʃkə</td>
<td>‘man, human’</td>
</tr>
<tr>
<td>hɛiʃ</td>
<td>hɛiʃkə</td>
<td>‘glove’</td>
</tr>
<tr>
<td>nɔːl</td>
<td>nɔːlɛkə</td>
<td>‘needle’</td>
</tr>
<tr>
<td>oːr</td>
<td>oːrkə</td>
<td>‘ear’</td>
</tr>
<tr>
<td>bɛl</td>
<td>bɛlkə</td>
<td>‘bell’</td>
</tr>
<tr>
<td>leɲc</td>
<td>leɲcə</td>
<td>‘ribbon’</td>
</tr>
<tr>
<td>ɛːɲ</td>
<td>ɛːɲkə</td>
<td>‘duck’</td>
</tr>
<tr>
<td>hɔut</td>
<td>hɔeycə</td>
<td>‘hat’</td>
</tr>
<tr>
<td>tant</td>
<td>tɛntə</td>
<td>‘aunt’</td>
</tr>
<tr>
<td>paçf</td>
<td>paçfkə</td>
<td>‘cap’</td>
</tr>
</tbody>
</table>
Distinctive features are capable of capturing the natural classes that can be identified in many languages. The theoretical claim they embody is that distributional generalizations and phonological processes arrange the segments they apply to in groups that possess some feature or feature combination. This suggests that phonetic processes which may initially affect specific segments in one generation become focused in a following generation on a group of segments that are defined by some feature specification. For instance, even though there are no words in English that begin with [spw] or [stw], we may predict that if they were introduced as loanwords, they would fall under the generalization that no aspiration or devoicing of approximants applies to segments after syllables beginning with [sc-], like sport or splint, regardless of whether late VOTs (see section 2.2.2) affected [w] in the donor language. Or again, while degeminization in Dutch applies to all consonants, it skips [−cons] [v], despite the fact that this consonant has a fairly firm contact between the upper teeth and the lower lip in Dutch as spoken north of the main rivers. The rival theory here is of course that all such groupings are a result of natural phonetic tendencies.

We have seen that segments may be unspecified for certain features, as a result of which natural classes are more straightforwardly expressed, such that generalizations may be simpler. For instance, a characterization that all fricatives except [f] are voiceless in some context loses its exception clause if only oral fricatives, i.e. not the glottal fricatives, have manner features like [+cont]. Finally, languages may apparently develop different specifications for the same phonetic segments, as suggested by the positive value of the feature [±continuant] in some languages and the negative one in other languages. This is an indication that distinctive features emerge during language acquisition rather than being innately listed.

NOTE

1 It has earlier also been assumed that [h?] are [+son], as in Chomsky and Halle (1968), Halle and Clements (1983) and Trommelen and Zonneveld (1983).
6.1 INTRODUCTION

In this chapter, we introduce the features that specify the place of articulation of consonants and the tongue position of vowels. Among the place features, there are four univalent features specifying the major areas of articulation. These are the features [labial], [coronal], [dorsal] and [radical]: either a segment has the feature, or it does not. This implies that, just as laryngeal segments are not specified for a number of features (those specifying either manner or place of articulation), not all segments will be specified for all the place features: a consonant which is not coronal will therefore not have the feature [coronal]. Within these major zones of articulation, binary place features are used to characterize the more detailed articulatory distinctions. The feature [±round], for example, will be used to specify segments that are articulated with the help of the lips, i.e. that are [labial]. Segments whose articulation does not involve any activity of the lips will thus be neither [+round] nor [−round]: they have no specification for either [labial] or [±round].

In this chapter we also present the notational devices Chomsky and Halle (1968) employed in their formulation of phonological rules, which often figure in phonological generalizations generally, whether these are expressed as rules or as constraints.

6.2 PLACE FEATURES

6.2.1 Labial

[labial] segments are articulated with the lips, like [f p m], or in the case of vowels are formed with lip rounding, like [y o æ]. Segments that are [labial] may be specified for [±round].

1. [±round]. [+round] segments have lip rounding, like [pʰ tʰ o u ɔ]; [−round] segments do not. In rare cases, unrounded and rounded labial segments contrast, as in Margi, Nambakaengo and Kilivila (Senft 1986). Labialized segments like [tʷ] will be discussed further in chapter 14.

6.2.2 Coronal

[coronal] segments are articulated with a raised crown of the tongue, i.e. a raised tip and/or blade, ranging from a dental [θ] to a prepalatal [j]. Examples of [coronal] segments are [t z ɬ ʃ j ɲ r]. [coronal] segments are further specified
for the features [±anterior] and [±distributed], and in the case of coronal fricatives and affricates also for [±strident].

1  [±anterior]. For [+ant] segments, the crown articulates with the alveolar ridge or somewhere further forward, while for [–ant] segments, the crown articulates with a point behind the alveolar ridge. Thus, [t d s z θ ð n l] are [+ant], while prepalatal or postalveolar and retroflex consonants (e.g. [c ʃ ʒ ɲ] and [t ɖ ʂ ʐ ɳ ɽ]) are [–ant].

2  [±distributed]. Segments that are [+distr] are produced with a constriction that extends for a relatively great distance along the vocal tract, while for [–distr] segments this is not the case. Thus, consonants produced with the tip of the tongue (apical consonants like British English [t d n]) are [–distr], as are [s z] (Clements 1985). Blade-articulated (laminal) consonants like [ʃ t ʒ] are [+distr]. Dental consonants like [θ ð t̪ l] are also [+distr], because even where it is only the tip that touches the front teeth, the blade is close to the alveolar ridge and in fact contributes to the acoustic effect. Retroflex consonants are [–distr]: the tip articulates with the part of the palate immediately behind the alveolar ridge. Australian languages frequently have a four-way opposition, utilizing the four possibilities given by these features (Butcher 2006). Four coronal stops and nasals contrast in Kayardild, for instance, as shown in (1) (Evans 1995).¹

<table>
<thead>
<tr>
<th></th>
<th>(Lamino-) dental</th>
<th>(Apico-) alveolar</th>
<th>(Lamino-) prepalatal</th>
<th>Retroflex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ant</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Distr</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Q48 In Bengali, there is an optional rule which deletes [r] before certain consonants, allowing the consonant to geminate. Characterize the class of consonants that trigger R-DELETION on the basis of the following data (Hayes and Lahiri 1991).

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>‘Meaning’</th>
</tr>
</thead>
<tbody>
<tr>
<td>barʃa</td>
<td>baʃʃa</td>
<td>‘rainy season’</td>
</tr>
<tr>
<td>mitʃu r ʃari</td>
<td>mitʃu ʃʃari</td>
<td>‘Mithu’s sari’</td>
</tr>
<tr>
<td>ram-er bari</td>
<td>ram-er bari</td>
<td>‘Ram’s house’</td>
</tr>
<tr>
<td>ram-er ʃaka</td>
<td>ram-et ʃa</td>
<td>‘Ram’s money’</td>
</tr>
<tr>
<td>fon-korbo</td>
<td>fon-korbo</td>
<td>‘will telephone’</td>
</tr>
<tr>
<td>ʃundor dɔrdʒa</td>
<td>ʃundod dɔddʒa</td>
<td>‘beautiful door’</td>
</tr>
<tr>
<td>bordi</td>
<td>boddi</td>
<td>‘elder sister’</td>
</tr>
<tr>
<td>bʃiorti</td>
<td>bʃiotti</td>
<td>‘full’</td>
</tr>
</tbody>
</table>
6.2.3 Dorsal

[DORSAL] sounds are articulated with bunched dorsum: [k g x ɣ ŋ] (velars), as well as [ç k] (fronted velars) and uvulars (e.g. [χ q]). In addition, all vowels are [DORSAL]. [DORSAL] segments are further specified for a set of features that specify just where the bunch of the tongue body is located, the **tongue body features**.

1. [+high]. Segments that are [+high] raise the dorsum to a position close to the roof of the mouth, while [−high] segments do not. Thus, [+high] segments are [i i ɣ y ŋ ʊ u], as well as [ç], and [k g x ɣ ŋ], while [χ e o a] are [−high], for instance.

2. [+low]. [+low] segments have the bunched dorsum low in the mouth, while [−low] segments do not. Thus, [+low] segments are [a ε ɔː], for instance.

3. [+back]. [+back] segments have the bunch of the tongue positioned in the centre or further back, while [−back] segments have the bunch in the front. Thus, [+back] segments are velar and uvular consonants (e.g. [k ɣ χ]) and vowels like [ʊ ŋ a o ɔ], while [−back] segments are fronted velars like [k] and [ç], and vowels like [i y o e]. Although [ç] is classed with the palatal consonants [c j n] in the IPA chart (i.e. with the [CORONAL] consonants), it is a fronted velar, i.e. a [DORSAL] consonant. Many languages have [ç] and [x] in complementary distribution depending on the backness of the preceding or following vowel. For instance, Greek [ˈçeri] 'hand' begins with the same phoneme as [ˈxari] 'charm'.

4. [+tense]. [+tense] vowels like [i e a u] are produced with a more peripheral and somewhat closer tongue position than their [−tense] counterparts [i e a u]. The feature is relevant only if the language has vocalic oppositions like [i – i], [y – y], [u – o], etc. It is commonly used in Germanic languages, which have contrasts like English [su:t] suit – [sot] soot and German [ˈmiːtə] ‘rental fee’ – [ˈmitə] ‘middle’. The features [+Advanced Tongue Root; ATR], as used, for instance, in the description of the West African language Akan (Lindau 1978), and [±Retracted Tongue Root; RTR], used, for instance, in the description of the Tungusic languages of Siberia (Li 1996), may be seen as phonetic variants of this phonological feature. [+ATR] involves a forward position of the tongue body, with concomitant enlargement of the pharynx, while [+RTR] involves a retraction and lowering of the tongue body, with concomitant narrowing of the pharynx. Akan has four plain vowels and five [+ATR] vowels, three of the former type occurring in [−ATR] [ɛbʊɔ] ‘stone’, and three of the latter
in [+ATR] [ebuo] ‘nest’. The Baiyinna variety of Oroch has nine plain and nine [+RTR] vowels, and [−RTR] [olo:] ‘to cook’, for instance, contrasts with [+RTR] [sla:] ‘to wade’. The features [±tense], [±ATR] and [±RTR], while phonetically somewhat different, appear never to co-occur in the same language (cf. Halle and Clements 1983: 7; Ladefoged and Maddieson 1996: 300).

6.2.4 Radical

[radical] (also [pharyngeal]) sounds are articulated with the root of the tongue. A voiceless fricative [h] occurs in many varieties of Arabic, as does a pharyngeal approximant [ʕ]. See Ladefoged and Maddieson (1996) for more information.

6.2.5 Some examples

Table 6.1 contains feature values for a number of representative consonants. Notice that many segments have more than one place feature and are technically ‘complex’. A round vowel like [u] is [labial, +round] as well as [dorsal, +back, +high], while a labialized [t] is both [coronal] and [labial]. In chapter 14 we will further discuss the representation of complex consonants.

A seven-vowel system like that of Italian comes out as in (2). It is illustrated in (3).

\[
\begin{array}{|c|c|c|}
\hline
\hline
i & u & \\
\hline
[−high, –low] & e & o \\
\hline
[−high, +low] & e & a & o \\
\hline
\end{array}
\]

(3)  
<table>
<thead>
<tr>
<th>milːe</th>
<th>‘thousand’</th>
<th>pupːo</th>
<th>‘fist’</th>
</tr>
</thead>
<tbody>
<tr>
<td>seta</td>
<td>‘silk’</td>
<td>sole</td>
<td>‘sun’</td>
</tr>
<tr>
<td>sempre</td>
<td>‘always’</td>
<td>skala</td>
<td>‘stairs’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fɔka</td>
<td>‘seal’</td>
</tr>
</tbody>
</table>

As observed above, in many Germanic languages, the vowels divide into a lax set and a tense set. Standard German has the vowel system in (4), where the lax vowel is given on the left of each cell. Examples are given in (5). The bracketed [æː:] has merged with [eː:] in the speech of many speakers. Note that [ɛ ɔ] are the lax counterparts of [eː oː] in German and Dutch.

\[
\begin{array}{|c|c|c|}
\hline
\hline
i i: & y y: & o u: & \\
\hline
[−high, –low] & e e: & ð o: & \\
\hline
[−high, +low] & æː & a aː & \\
\hline
\end{array}
\]
Table 6.1
Feature specifications of 23 representative consonants for 16 features. Binary features are specified as + or −, while the presence of a unary feature is indicated by √. Blanks indicate that the consonant is not specified for the feature. The value of the feature [continuant] for [t] varies across languages.

<table>
<thead>
<tr>
<th>Feature</th>
<th>cons</th>
<th>son</th>
<th>approx</th>
<th>cont</th>
<th>nas</th>
<th>lat</th>
<th>voice</th>
<th>spread</th>
<th>LABIAL</th>
<th>COR</th>
<th>distr</th>
<th>ant</th>
<th>strid</th>
<th>DORSAL</th>
<th>high</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>t</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>pʃ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>tʃ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>kʰ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>b</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>f</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>θ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>s</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>ζ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>γ</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>v</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>y</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>m</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>&amp;</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>η</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>η</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>r</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>l</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>j</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>w</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>h</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Historically, the Germanic tense-lax contrast arose from a quantity distinction, as evidenced by the correlation between the features long and tense in German. However, in many varieties of West Germanic they are independent, such that lax vowels may be long, and tense vowels short. Standard Dutch has short tense [+high] vowels, for instance, as seen in (6).

There are languages that contrast four vowel heights without employing a tense-lax contrast, of which Danish and Imonda are examples. The tongue-height features [+high] and [±low] cannot characterize such systems. Of the four theoretically possible combinations, the specification *[+high, +low] must be ruled out, because it is contradictory: the body of the tongue cannot simultaneously be raised and lowered. The Imonda vowel system is given in (7) (Seiler 1985). Clements and Hume (1995) present a feature framework that can account for such four-height systems (as well as for five-height systems, which have also been reported to exist).

Q49 In Corsican, some vowels are nasalized before a nasal consonant in the same syllable (Agostini 1995). How would you characterize the class of vowels that undergo this nasalization?

| 'printʃile | 'prince' |
| 'áŋku | 'also' |
| 'pɒnte | 'bridge' |
Q50 Yanyuwa has four contrastive coronal places of articulation, illustrated below for voiced plosives. (The lamino-postalveolar one is like [dʒ] without the fricative release.)

<table>
<thead>
<tr>
<th>Lamino-dental</th>
<th>Apico-alveolar</th>
<th>Apico-postalveolar</th>
<th>Lamino-postalveolar</th>
</tr>
</thead>
<tbody>
<tr>
<td>d̪</td>
<td>d</td>
<td>d̠</td>
<td>d</td>
</tr>
</tbody>
</table>

1. Show how the features [±anterior] and [±distributed] can characterize each of these articulation places.
2. Yanyuwa has a further contrast between a voiced palatal and velar plosive [i] and [g]. Assuming these are both [dorsal], how would you characterize these places of articulation?

### 6.3 REDUNDANT VS CONTRASTIVE FEATURES

Frequently, feature specifications are predictable. This predictability is in part a consequence of the incompatibility of particular feature specifications. For instance, as we have seen in the previous section, the presence of [+low] in the specification of a vowel predicts that it will be [−high]: the tongue cannot be raised and lowered at the same time. The same goes for [+high]: it predicts [−low]. In large measure, however, the predictability is language-specific and results from the fact that not all languages use the phonological possibilities to the full. Take the Turkish vowel system, for instance. This language employs all the combinations of [±back] and [±round], but has only two distinctive vowel heights, as shown in (8), with examples in (9).

(8)

<table>
<thead>
<tr>
<th></th>
<th>[−back, −round]</th>
<th>[−back, +round]</th>
<th>[+back, −round]</th>
<th>[+back, +round]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+high]</td>
<td>i</td>
<td>y</td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td>[−high]</td>
<td>e</td>
<td>ø</td>
<td>a</td>
<td>o</td>
</tr>
</tbody>
</table>

(9)        

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>dif</td>
<td>‘tooth’</td>
<td>gyl</td>
<td>‘rose’</td>
</tr>
<tr>
<td>kep</td>
<td>‘cap’</td>
<td>gol</td>
<td>‘lake’</td>
</tr>
</tbody>
</table>

In this system the feature [±low] is entirely redundant, as one of its values, [+low], does not appear to be used at all. In other inventories, both values of a feature may occur, but one value may be predictable from other features. In the
Describing places of articulation

Italian vowel system (2), for instance, [+round] predicts [+back]: all round vowels in this language are back. (This absence of front rounded vowels, incidentally, is a characteristic that Italian has in common with most languages in the world.) Or again, while in many languages obstruents come in pairs contrasting for [+voice], like [p b], [s z], etc., sonant segments like [m l r] are usually always voiced, with the result that in those languages [+son] predicts [+voice]. This means that while in the class of obstruents the feature [+voice] is contrastive (or distinctive), [+voice] is redundant in sonorants: they are predictably [+voice].

A large body of work, often referred to as **Underspecification Theory**, has been devoted to the question of whether redundant features should be included in the underlying representation of morphemes, and, if not, when they should be supplied. Just as in the case of the non-specification of manner and place features in glottal consonants (see section 5.7.2), arguments for or against underspecification have been based on the simplifying effect that the absence of a feature has on the formulation of phonological generalizations. We give one example here, from Steriade (1987). If we assume that [l] is [+cont] in Latin, [+lateral] is distinctive within the class of liquids, [l r], or [+approx, +cons]. In the class of nasals and obstruents ([-approx]) as well as in the class of glides and vowels ([−cons]), the feature is redundant. Steriade (1987) proposes that only contrastive features are specified in underlying representations. This means that [r] is marked as [−lat] and [l] as [+lat], but that in no other segment is there a specification for [±lateral]. This assumption allows for an interesting description of the distribution of the alternants -aris and -alis of the adjectival suffix. Consider the forms in (10).

(10) a nav-alis ‘naval’
    crimin-alis ‘criminal’

b sol-aris ‘solar’
    milit-aris ‘military’

c flor-alis ‘floral’

The forms in (10a,b) suggest that the alternation depends on the presence of the feature [+lat] in the base: if the base contains [+lat], the form -aris is used, otherwise -alis. If we assume the underlying form of the suffix to be /alis/, the [+lat] [l] is seen to change into [−lat] [r] after a base containing [+lat]. (Such processes are known as **dissimilations**.) But now consider (9c). Although the base contains [+lat], due to the presence of [l] in [fl-], the suffix is alis. It is here that the underlying [−lat] specification of [r] comes in useful. Because other nonlateral sounds, like [t] of militaris, are not specified as [−lat], Steriade’s proposal allows us to say that the last specification for [lat] in the base determines the specification for [lat] of the first consonant in the suffix. This is shown in (11).

(11) [+lat] [−lat]

a mil
    [+lat] [−lat] [−lat]
    [+lat]

b fl
    [−lat] [−lat] [−lat] [−lat]
    or-a
    [−lat]
    l
    is
This particular conception of underspecification, in which features are specified only in segments that contrast for the feature concerned, is known as **contrastive underspecification**. The topic of underspecification is closely related to one's view of how segments are represented. A rival view is that underlying representations only ever include one of the two feature specifications of contrastive features.

**Q51** In Luganda, [r] and [l] occur in complementary distribution (Chess-was 1963).

1. What determines their distribution? List the contexts in which each allophone occurs.
2. Which of these two contexts is statable in terms of distinctive features?
3. Which of the segments would you choose as the underlying one? Please motivate your answer.

**Q52** Use minimal numbers of distinctive features to characterize the five natural segment classes within the segment inventory of Telugu.

<table>
<thead>
<tr>
<th>p</th>
<th>pʰ</th>
<th>t</th>
<th>tʰ</th>
<th>ʈʃ</th>
<th>t̪</th>
<th>t̪ʰ</th>
<th>k</th>
<th>kʰ</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>bʰ</td>
<td>d</td>
<td>dʰ</td>
<td>ḍ</td>
<td>ḍʰ</td>
<td>g</td>
<td>gʱ</td>
<td></td>
</tr>
<tr>
<td>s</td>
<td>s̈</td>
<td>m</td>
<td>n</td>
<td>n̄</td>
<td>l</td>
<td>r</td>
<td>ɾ</td>
<td>l̥</td>
</tr>
</tbody>
</table>

1. pʰ bʰ tʰ dʰ ʈʰ ḍʰ kʰ gʱ
2. t tʰ d dʰ t̪ t̪ʰ ḍ ḍʰ
3. b bʰ m
4. w j i e a o u
5. a o
Q53 Use minimal numbers of distinctive features to characterize the five natural segment classes within the segment inventory of Amharic (Hayward and Hayward 1992).

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t’</td>
<td>t&quot;</td>
<td>c</td>
<td>c’</td>
<td>k</td>
</tr>
<tr>
<td>b</td>
<td>b’</td>
<td>d</td>
<td>j</td>
<td>g</td>
<td>g’</td>
</tr>
<tr>
<td>f</td>
<td>f’</td>
<td>s</td>
<td>s’</td>
<td>ʃ</td>
<td>h</td>
</tr>
<tr>
<td>z</td>
<td>ʒ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>m’</td>
<td>n</td>
<td>ɲ</td>
<td>ɲ</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i</td>
<td>u</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>ɪ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>f</td>
<td>f’</td>
<td>s</td>
<td>s’</td>
<td>ʒ</td>
</tr>
<tr>
<td>2</td>
<td>t’</td>
<td>t&quot;</td>
<td>c’</td>
<td>k’</td>
<td>s’</td>
</tr>
<tr>
<td>3</td>
<td>c</td>
<td>c’</td>
<td>ʃ</td>
<td>ʃ</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>n</td>
<td>ɲ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>u</td>
<td>o</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6.4 WRITING RULES

Phonological rules are formal expressions that describe changes in the phonological representations of words. As a result of the application of a phonological rule, a segment may be inserted or deleted, or one or more of its feature values may be changed (Chomsky and Halle 1968, SPE). It will also be shown that the assumption that rules are ordered, whereby the output of one rule serves as the input to the next, allows for simple phonological generalizations.

In the representation proposed by SPE, a segment is a list of feature specifications, also referred to as a feature matrix. Although in the original proposal all features were binary, we give the representation with the univalent place feature [CORONAL] in (12b). An expression like ‘[t]’ is therefore a shorthand notation for the corresponding feature matrix.

A morpheme is represented as a string of feature matrices. The beginnings and ends of morphemes are indicated by boundary symbols. What may, with hindsight, be seen as puzzling is that the only boundaries were morphosyntactic. Syllableity was handled by a feature, [−syl] or C vs [+syl] or C. The symbol + was included in representations at the boundaries of word-internal morphemes (morpheme boundary), while # served as the word boundary. Thus, (13) is the representation of the word ‘pens’. In much of the literature, word-internal morpheme boundaries are represented by means of a dash, as in [pen-z] ‘pens’, rather than with +.
Sentences are represented as strings of words. The position in SPE was one in which the morphology and syntax preceded the phonology: only when the words are inserted into the sentence will the phonological rules be called upon to make the necessary adjustments. In (14), a 'shorthand' representation is given of the sentence *Pens leak*.

(14)  

6.4.1 The general *SPE* rule format

The general format of an *SPE* rule is as given in (15).
Describing places of articulation

\[ (15) \quad \begin{bmatrix} \text{Minimal feature specification of segment(s)} \\
\text{undergoing the change} \end{bmatrix} \rightarrow \begin{bmatrix} \text{Features that change} \end{bmatrix}, \quad \begin{bmatrix} \text{Minimal specification of lefthand context} \\
\text{Minimal specification of righthand context} \end{bmatrix} \]

The information to the left of the arrow is the **focus** of the change, that between the arrow and the slash is the **structural change** (SC), while the information to the right of the slash is the **context**. The focus plus the context is known as the **structural description** (SD). Rules are assumed to apply within words, and if a rule is to apply across word boundaries, the SD needs to include the symbol #. In contrast, rules are assumed to apply across +-boundaries, even if no + is specified. (If a rule applies only at a +-boundary, this must be included in the SD.) To illustrate, we give (16) as the rule that devoices voiced obstruents at the word end, as occurring in German and Dutch, for instance.

\[ (16) \quad \text{final devoicing} \quad [-\text{son}] \rightarrow \quad [-\text{voice}] / \_ \_ \_ # \]

The rule scans the feature matrices from left to right for the presence of [-son]; then, if it finds one, it checks whether there is a # on the right and, if there is, it specifies ‘-’ as the value of the feature [voice] in the matrix with [-son]. In many cases, the obstruent will already be [-voice]. The rule is then said to apply vacuously: the SD is satisfied, but application of the rule does not bring about a difference. The reason for doing it this way is that the rule can be written with fewer terms than if we were to add the feature [+voice] to the focus of the rule.

The obstruent undergoing rule (16) frequently appears in word-final position. However, sometimes an obstruent needs to be devoiced that is not adjacent to the word boundary. For instance, when the 3sg suffix [t] is attached to the verb stem [le:z] 'read', the surface result is not *[le:zt] but [le:st]. In the SPE representations, this means that we must express in the rule that other obstruents may intervene between the obstruent to be devoiced and the word end. This is done by adding the term ‘C\_\_t’ before the #. The C itself is shorthand for a segment outside the nucleus of the syllable, while V is similarly used to mean a vowel in the syllable nucleus. The subscripted number n means ‘n or more’ (instances of the symbol), just as a superscripted digit means ‘n or less’. The term ‘C\_\_t’ therefore translates as ‘zero or more consonants’ and has the effect of allowing rules to ignore consonants in the positions in which it is used. Our rule therefore now looks like (17).

\[ (17) \quad \text{final devoicing} \quad [-\text{son}] \rightarrow \quad [-\text{voice}] / \_ \_ \_ C\_t # \]

Version (17) would be able to apply twice in the same word. When shifting through the form from left to right, the rule will find that its SD is in fact met in a form like [le:z-t] 'read-3sg'. Similarly, it is met twice in a case like [ho:vd] 'head'.
whose singular and plural forms are [hoːft] and ['ho:v.dɔn], respectively. The first focus is [v], when [d] corresponds to C₀, and the second is [d], when C₀ corresponds to no consonant.

Q54 The Dutch past participle is formed by prefixing [χə] and suffixing [d] to the verb stem. The suffix shows up in its underlying form when the inflectional suffix [ə] is added, as in [χə+vyl+d+ə] ‘PARTIC+fill+PARTIC+INF’, but otherwise shows up as [t]. In the case of a verb stem like [sχyld], the uninflected [χə+sχyld+d] may be formed, and is pronounced [χəsχyl]. Does rule (17) apply to both underlying [d]s in this last form? Motivate your answer.

In retrospect, the notation ‘C₀’ was really a way of avoiding reference to syllables. The devoicing rule really applies to the syllable coda, as shown by the way speakers of Dutch treat English words like Sidney [ˈsdni], business [ˈbɪznɪs]. In loans, English syllable-final [d ə] are replaced with [t s], respectively (Booij 1995). Venneman (1972), Hooper (1976) and Kahn (1976) (re)introduced the syllable in phonological representation. Reference to the syllable and its constituents makes it possible to give a formulation of final devoicing which takes these additional facts into account. In (18), the σ-labelled parenthesis is used to indicate membership in the syllable coda. In chapter 9 we will deal more extensively with the syllable in the structural descriptions of rules.

(18) final devoicing [−son]  →  [−voice] / _ C₀

6.4.2 Additional notations

A notational device that has been widely discredited as a theoretical element, but that is often used out of convenience, is the brace. The brace notation is used to express a disjunction between two or more terms (‘either . . . or’), and is thus found in rules that are partly identical. Again, when no reference to the syllable is made, the syllable-final context can often be captured by saying ‘before a consonant or the word end’. Rule (19), for instance, nasalizes vowels before a nasal followed by another consonant or the word end. It would change French [bɔn] ‘good-masc’ and [bɔnte] ‘goodness’ into [bɔ̃n] and [bɔ̃nte], respectively.

(19) V  →  [+nas] / _ [+nas] {C} #

Q55 The empty-set symbol Ø is used to the left of the arrow in the case of insertion, and to the right of the arrow in the case of deletion. What would be the prose version of (1), for instance? And of (2)?

(1) Ø  →  ə /V_.r#
(2) r  →  Ø/_.{C} #
To express assimilations, feature values need to be made to agree. For instance, in Turkish, vowel harmony requires that high vowels in suffixes agree for [back] and [round] with the preceding vowel in the word, as shown in (20).

(20)  
<table>
<thead>
<tr>
<th>Nominative</th>
<th>Possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>köj</td>
</tr>
<tr>
<td>b</td>
<td>kep</td>
</tr>
<tr>
<td>c</td>
<td>at</td>
</tr>
<tr>
<td>d</td>
<td>son</td>
</tr>
</tbody>
</table>

A rule to achieve this would use variable feature values, expressed with the help of Greek letters. Rule (21) says that high vowels agree in backness and roundness with the preceding vowel: α could be − or +, and so could β, independently.

(21) $V^{+\text{high}} \rightarrow [\alpha_{\text{back}} / \beta_{\text{round}}] C_0 + C_0$

If the target vowel in the suffix is [−high], only the backness features in the stem and suffix agree. To express such conditional dependence, angled brackets have been used. In (22), this further specification has been indicated: all target vowels agree in backness with the last vowel of the stem, and if the target vowel is [+high], it also agrees in roundness with the last vowel of the stem.

(22) TURKISH VOWEL HARMONY $V^{<+\text{high}>} \rightarrow [\alpha_{\text{back}} / \beta_{\text{round}}] C_0 + C_0$

Q56

1 Write out the four readings of (21).
2 The plural suffix is either [ler] or [lar]. Which items in (20) take [lar]?

The parenthesis notation includes optional elements in rules. Dutch has a rule of regressive voicing which applies within words as well as across word boundaries. Within words, it applies to [lif-daemon] ‘dear-ness, love’ to form [livda], and across words it applies in [lif #di:ran] ‘dear animal’ to form [liv di:ran]. Since SPE assumes that every word that leaves the morphology has #s around it, two #s need to be specified in the SD. However, because the rule also applies within words, they need to be put in parentheses to indicate that they may, but need not, be present in the representation, as shown in (23).

(23) [−son] $\rightarrow$ [+voice] / ∅ (#{}) − cont − son

+ voice

− cont

− son
There are processes that affect more than one segment. For example, metathesis is a process that switches round two segments, as can be seen in the Old English word for grass, which varied between [græs] and [gærs]. To be able to refer to changes involving more than one segment, the transformational rule format was used. It lists the relevant string of segments and boundaries to the left of the arrow, and repeats that string, with the SC, to the right of the arrow. The segments of the context are not literally reproduced, but identified with the help of digits. Rule (24) says: ‘Delete a coronal nasal before a consonant or at the word end, and nasalize the vowel that precedes it’. It would change French [bon] and [bonte] into [b3] and [bôte], respectively.

\[
\begin{align*}
\text{(24)} & \quad [V] \quad \rightarrow \quad \# \quad \rightarrow \quad 1 \quad 3 \\
\text{V} & \quad -\text{cons} \quad +\text{nas} \\
\text{COR} & \quad \rightarrow \quad 0 \\
1 & \quad 2 \quad 3 \\
\end{align*}
\]

Q57 Assuming the segment inventory given below, give prose statements of rules (1) to (3).

\[
\begin{align*}
\text{p} & \quad \text{t} \quad \text{c} \quad \text{k} \quad \text{i} \quad \text{u} \\
\text{b} & \quad \text{d} \quad \text{j} \quad \text{g} \quad \text{e} \quad \text{o} \\
\text{f} & \quad \text{s} \quad \text{ʃ} \quad \chi \quad \text{a} \\
\text{v} & \quad \text{z} \quad \text{ʒ} \\
\text{m} & \quad \text{n} \quad \text{ɲ} \quad \text{ŋ} \\
\text{l} & \quad \text{r} \\
\end{align*}
\]

(1) \[
\begin{align*}
\text{V} & \quad \rightarrow \quad [+\text{high}] / _-\text{#} \\
\text{low} & \\
\end{align*}
\]

(2) \[
\begin{align*}
\text{V} & \quad \rightarrow \quad _-\text{low} / _-\text{high} \\
\text{back} & \\
\alpha & \quad \beta \\
\end{align*}
\]

(3) \[
\begin{align*}
\text{Ø} & \quad \rightarrow \quad _-\text{cont} / _-\text{voice} \\
\text{COR} & \quad _-\text{nas} / _-\text{ant} \\
\alpha & \quad \beta \\
\end{align*}
\]
Q58 Assuming the segment inventory given in the previous question, write formal rules:

1. Obstruents are voiceless after word-internal or word-external obstruents.
2. An [ə] is inserted between [r l] and a labial obstruent in the same word.
3. Suffix-initial vowels are front if the preceding syllable contains [i], and back if the preceding syllable contains [u].

6.5 LINEAR ORDER

Obviously, if in some language there were only one rule, its input would consist of the lexical representation of the forms that meet its structural description (the underlying representation, UR), while its output would correspond to the actual pronunciation (the surface representation, SR). Languages generally have more than one rule, and the question therefore arises of how phonological rules apply in a rule-based analysis: in sequence or simultaneously? Let us assume, for the sake of argument, that rules apply simultaneously to the underlying representation. Any rule would always scan the underlying representation to see if its structural description was met and, if it was, the change would be made. This mode of application is known as simultaneous rule ordering. Another possibility would be that the output of one rule is taken as the input of another rule. In this situation, you would have to know in which order the rules applied. This option is known as linear rule ordering. A brief consideration of simultaneous rule ordering will lead to a rejection of that option. Take n-deletion and ə-insertion in Dutch. The first deletes word-final [n] after [ə], while the second inserts [ə] between a liquid and a consonant other than [t d]. In (25), the underlying forms of the first two words, ‘calf’ and ‘bedsheet’, undergo one of these rules only, but the third form, ‘fern’, is a potential target for both. Simultaneous application would prevent n-deletion from applying, as there is no schwa to its left. To get the output form [vaːrə], which is also common, we must first apply ə-insertion so as to create the intermediate form [vaːrən]. Now n-deletion can apply so as to give the alternative form, as in (26).

To salvage the hypothesis that rules apply simultaneously, we would have to provide n-deletion with an enriched SD: ‘after schwa or [r]’. Not only does this carry
the disadvantage of an otherwise unnecessary addition to the context, it would also be impossible to capture schwa and [r] in a single feature specification without special assumptions about the feature matrices of those two segments. By insisting on serial rule ordering, the rule-based theory of Chomsky and Halle (1968) also achieved that varieties of the same language can be described by reversing rule orders. As it happens, [vaːrən], the output of (25), is the more usual pronunciation of that word.

As one further example of how, in a rule-based description, linear order avoids complicated phonological generalizations, take the two rules of English which apply to the plurals of nouns. As you will remember, the plural suffix has three alternants. After [s z ʃ ʒ tʃ dʒ], the alternant [iz] is used; after voiceless segments other than [s ʃ tʃ], [s] occurs, while elsewhere [z] is used, which we argued in chapter 4 is also the underlying form of the morpheme. The plurals of [bæz], [pen] and [bok] (bus, pen, book) are, respectively, [bæsz], [penz] and [boks]. We thus need one rule that inserts [1] between the stem and the suffix if the stem ends in [s z ʃ ʒ] (1-INSERTION), and one that devoices [z] after voiceless segments (DEVOICING). We repeat this derivation in (27).

(27)  

<table>
<thead>
<tr>
<th></th>
<th>bæz-z</th>
<th>pen-z</th>
<th>bok-z</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-INSERTION</td>
<td>bæsz</td>
<td>(n.a.)</td>
<td>(n.a.)</td>
</tr>
<tr>
<td>DEVOICING</td>
<td>(n.a.)</td>
<td>(n.a.)</td>
<td>buks</td>
</tr>
<tr>
<td></td>
<td>bæsz</td>
<td>penz</td>
<td>buks</td>
</tr>
</tbody>
</table>

Which theory allows us to write the simpler rules, the simultaneous-ordering theory or the linear-ordering theory? For the formulation of 1-INSERTION, given in (28), it would not make any difference which theory we adopted, because under either theory the rule applies to the underlying form. This is necessarily so under a simultaneous mode of application, and in an ordered mode of application, it would simply be ordered first.

(28)  

Now consider the second rule, DEVOICING. In a simultaneous-ordering theory, we would have to make sure that it does not devoice the [z] in cases like [bæsz]. That is, the rule needs to express ‘devoice [z] after all voiceless segments except [s ʃ tʃ], i.e. after [p t k f θ]. However, [p t k f θ] do not form a natural class: phonological rules do not typically refer to this group of sounds, forcing us to use the suspect ‘either . . . or’ braces. The formulation in (29) says: ‘after a voiceless obstruent which is either [−strid] (i.e. [θ]) or [labial] (i.e. [p f]) or [−cont] (i.e. [p t k]).’ To make matters worse, the feature [−cont] includes the affricate [tʃ], so that (29) still makes the wrong prediction that [z] devoices in words like batches ([bætʃiz], from [bætʃ-z]). In order to single out [p t k] to the exclusion of [tʃ] we could replace [−cont] with the feature combination [−cont, −strid], since only the affricate is [+strid].
Now consider the other option. If devoicing is allowed to apply to the output of i-insertion, we could simplify (29) to (30). This is because at the time that devoicing gets to apply to [bʌz] (the output of i-insertion), the [z] no longer appears after a voiceless segment, but after [+voice] [t].

### Conclusion

Segments can be analysed into collections of univalent and binary features. A successful feature analysis will characterize the segmental contrasts that are used in the languages of the world, while it will additionally allow natural feature classes, groups of segments that figure in phonological grammars, to be referred to. We have seen that segments may be unspecified for certain features, because the feature is irrelevant for the class of segments concerned, either universally, like [±strident] in the case of vowels or [±lateral] in the case of [h], or in some language, like [±low] in the case of the vowels of Turkish. Univalent features will be absent in all segments to which they do not apply, like [LABIAL] in the case of [t]. We have also made a distinction between contrastive and redundant features, where the former are features that, in some class of segments in some language, characterize contrasts, like [±voice] in a language with voiced and voiceless obstruents, and the latter are not, like [±voice] in a language that only has voiceless obstruents and voiced sonorant consonants and vowels. It is often assumed that redundant features do not appear in underlying representations.

### Note

1 This predicts that no language contrasts dental stops, which are [+ant, +distr], with laminaly produced alveolar stops, which are also [+ant, +distr], which would seem to be correct.
7.1 INTRODUCTION

The variation in the pronunciation of words is truly mind-boggling. Biological differences between speakers lead to different acoustic outputs, and independently of whether they are adult or child and male or female, people speaking the same language will have different accents, depending on their social class and the region they grew up in. Within the speech of the same speaker, the pronunciation of words will vary with the degree of formality. And even if we tried to remove all variation from our speech, we would not be able to produce acoustically identical pronunciations of the same word, due to the inevitable variation in the physiological mechanics and physical conditions of our environment. In this chapter, we will consider in more detail the type of within-speaker variation which is central to our topic, that which is a function of the phonological context. In the English spoken in North America, the final consonant in an expression like Right! will be an unreleased [t’], but in Right on!, this same /t/ will be an alveolar flap [ɾ]. The pronunciation of the English past tense suffix is either [t] or [d], depending on the voicing of the preceding segment, giving [lʊkt] as the past tense of look [lʊk], but [bɛgd] as the past tense of beg [bɛɡ]. And if the verb stem ends in [t] or [d], we find it is pronounced [ɪd], as in [ˈbaʃɑːtɪd] and [ˈniːdɪd], the past tense forms of buffet [ˈbaʃət] and need [niːd]. Variation in a segment’s pronunciation which is determined by the phonological context is known as ‘allophony’, or allophonic variation. It is discussed in section 7.2, where it is distinguished from stylistic variation, which is explained by the degree of formality of the speech.

Phonologists have responded to allophony and the segmental variation in the pronunciation of morphemes of the kind illustrated for the English past tense suffix by assuming multiple levels of representation. In the case of allophony, we are normally dealing with ‘novel’ sounds, like [tʰ] and [t’], which do not figure in underlying representations. In the mental lexicon of a speaker of American English, unreleased final plosives appear in the same shape and form as do aspirated plosives in word-initial position. That is, tight is listed as [tæt], not as [tʰæt’]. Importantly, too, the existence of two levels of representation will equally make it possible to express the phenomenon that a morpheme has a single phonological form at some cognitive level of representation, but has a number of phonological forms in the surface pronunciation, the morpheme alternants of chapter 4. Here, the surface morpheme alternants are represented in terms of segments that also exist in underlying
representations. For instance, the surface [s], [z] and [rz] of the regular English plural suffix also appear in underlying forms like *hiss*, *fez* and the pre-final syllable in *socialism*. It also happens that what are different sounds underlyingly show up as the same sound on the surface, a phenomenon known as **neutralization**.

### 7.2 ALLOPHONIC AND STYLISTIC VARIATION

The ‘segments’ referred to so far in this book typically display a good deal of variation, even in the speech of a single speaker. Two factors are largely responsible for this intraspeaker variation. One is **style**, or the degree of formality of the speech situation. An utterance like *Right!* will typically have an unreleased final plosive [t'], as explained in section 7.1, but in formal speech styles the [t] may well be released (so that you can hear a weakish [s]-like sound after it). In sociolinguistic studies, the **stylistic variants** [t'] and [t] are accordingly seen as the possible values of a phonological variable (t), whose frequencies of occurrence are compared across different speech styles for different groups of speakers. The second type of variation is due to variation in the **phonological context**. Such context-dependent variants are called **allophones**. The term **phoneme** is used to refer to the segment category to which the various allophones and stylistic variants belong. To distinguish the phoneme as a segment category from the individual allophones, it is often placed between slashes: / /. For example, in British English, there exist two rather different pronunciations of the phoneme /l/. When it occurs in the onset, it is pronounced as an alveolar lateral approximant, perhaps with some slight palatalization, as in *leek*, *follow*, a segment which is known as ‘clear l’. In the coda, as in *ill*, *cold*, its alveolar contact is accompanied by retraction of the tongue body towards the uvula, a sound symbolized by [l] and commonly known as ‘dark l’. Again, in many languages, coronal consonants alternate with palatoalveolar or alveolopalatal consonants, whereby the palatoalveolar one appears before [i], or sometimes before other front vowels as well. Thus, in Korean, underlying [sip'sam] is pronounced [cipsam] ‘thirteen’, while Igbo has [ʃ] before [i e] and [s] elsewhere, as shown by [ʃjiri] ‘he cooked’, [ʃjere] ‘he said’ vs [osere] ‘he wrote’ and [osara] ‘he washed’ (Jones 1967: 21). These examples illustrate the two kinds of phonological contexts that determine allophony. One is a particular segment or segment group, and the other is a structural position like the coda or the onset of the syllable.

Although allophonic variation in different languages may show some similarity, as suggested by the example of [s] before front vowels, it is important to see that languages will differ in the segmental differences they use to differentiate between words. That is, a given segmental difference may be phonemic in one language and allophonic in another. While the difference between [s] and [ʃ] is allophonic in Igbo, being fully conditioned by whether [i e] or one of the other vowels of the language follows in the word, in English the difference is **contrastive**: it is used to differentiate between morphemes. When a segmental difference is contrastive, there will typically be pairs of words that are distinguished only in that one has one
Underlying and surface representations

segment where the other has the other segment. Examples of such **minimal pairs** in English are [sɪp – ŋɪp] (sip – ship) and [liːs – liːʃ] (lease – leash).

**Q59** For each of the three following languages, say whether the difference between [r] and [ɾ] is (a) allophonic, (b) phonemic or (c) stylistic. Motivate your answers.

1. In some types of Southern Swedish, [ɾ] always appears at the beginning of a syllable and [r] always at the end.
2. In some types of Dutch, [r] is the usual realization in words like [proˈʁama:] ‘programme’. Some speakers of those varieties use this pronunciation in everyday life but use [ɾ], as in [proˈʁama], when announcing programmes on national television.
3. In Provençal, the word for ‘evening’ is [sɛro] and the word for ‘saw (noun)’ is [sɛro] (Jones 1967).

Allophonic variation is entirely predictable. As such, there is no reason to supply allophonic variation in the lexicon. The only lateral segment to appear in the lexicon of British English is just plain [l]. In a rule-based description, a rule of l-*velarization* would be postulated that adds the velarization to the specification of this consonant in the context ‘end of syllable’. This is shown in (1) for the words *leak* and *ill*.

(1) Underlying liːk il
   1-VELARIZATION (n.a.) ɫ
   Output liːk ɪɫ

Allophonic differences typically arise because particular contexts invite particular adaptations in the production of the segment. As a result, allophones are in **complementary distribution**. The phonological context in which [ʃ] occurs in Igbo (before [i e]) is the complement of the phonological context in which [s] occurs (before the remaining vowels of the language): together they make up the phonological context in which the underlying segment [s] occurs.

**Q60** In Tolitoli, an alveolar lateral approximant [l] is in complementary distribution with a retroflex lateral flap, given as [l]. What determines their distribution (Himmelmann 1991)?

<table>
<thead>
<tr>
<th>Word (Tolitoli)</th>
<th>Word (Provençal)</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>molo.to</td>
<td>mologo</td>
<td>‘wash hands’</td>
</tr>
<tr>
<td>too.to.l</td>
<td>tolitoli</td>
<td>‘Tolitoli’</td>
</tr>
<tr>
<td>u.lag</td>
<td>ulag</td>
<td>‘snake’</td>
</tr>
<tr>
<td>kiki.lo</td>
<td>kikilo</td>
<td>‘firefly’</td>
</tr>
<tr>
<td>membe.mbela</td>
<td>membembelan</td>
<td>‘to tremble’</td>
</tr>
<tr>
<td>mongiu.lan</td>
<td>mongiu lan</td>
<td>‘to choke’</td>
</tr>
<tr>
<td>lelem.batlan</td>
<td>lelembalan</td>
<td>‘to carry on the shoulder with a stick’</td>
</tr>
</tbody>
</table>
A complementary distribution of segments can also arise in a different way, such as when one segment happens to be restricted to the onset position and another to the coda position. In such cases, the segments are not historically related in the way that contextual allophones like [s] and [ʃ] are. It will then depend on the similarity of the two segments whether the phonology treats them as contextual variants of the same underlying segment. A well-known case is presented by [ŋ] and [h] in English. Because [h] may not occur at the end of a syllable and [ŋ] may not occur at the beginning of a syllable, the two segments are in complementary distribution. (That is, [hæŋ] ‘hang’ is a word, but neither *[tæh] nor *[ŋæt] could be a word.) There is general agreement that [h] and [ŋ] are underlying segments, since there is no plausible way in which one can be derived from the other.

Q63 French words like réseau [rezɔ], fraises [fʁɛz] and zone [zon] tend to be pronounced as [reso, fres, son], respectively, by Spanish speakers.
Similarly, they tend to pronounce French poison [pwazɔ̃] ‘poison’ and poissons [wasɔ̃] ‘fish’ both with [s]. Even though Spanish doesn’t contrast [s] and [z], [z] occurs in words like desdén [dez’den] ‘disdain’, mismo [’mizmo] ‘same’ and más vueltas [maz ’bweltas] ‘more laps’.

1. Are [s] and [z] different phonemes in Spanish? Motivate your answer.

2. How would an untrained Spanish speaker of French pronounce se laisse désirer [so ləs dezire] ‘is very desirable’: as [se les desire], [se lez desire], [se les dezire] or [se lez dezire]? Motivate your answer.

7.3 TWO LEVELS OF REPRESENTATION

Three general arguments have been advanced for the assumption of different levels of representation.

1. One argument is economy. Why supply allophonic information in the lexical entries if it can be stated in a set of allophonic rules that are valid for all morphemes of the lexicon? The strength of this argument has been called into question by Kenstowicz (1994a: 69), who points out that it is not self-evident that the descriptive economy achieved by having allophonic information supplied by rules should be reflected in the actual phonological representations in the mental lexicon. There is apparently no reason to assume that the brain could not store all that redundant information for each word in which it occurs. However, the issue here is not just storage capacity, but also, or even mainly, search time in speech perception (Lahiri and Marslen-Wilson 1991). It may be expected to be more difficult to retrieve the correct phonological form from a set of fully specified, hence complex, representations than from a set of more economical representations.

2. A second argument for the assumption of two levels of representation is that with a single level it would not be possible to express the phonological relatedness of morpheme alternants. We take the English plural suffix as an example. Suppose that instead of saying that there is a morpheme [z] which in different phonological contexts is adjusted in order to obey the phonological constraints of English, we were to say that three allomorphs are listed in the lexicon, [z], [s] and [iz], each of which is used in a specific phonological environment: [iz] after sibilants, [s] after (other) voiceless segments and [z] in other cases. This description of the regular plural formation would be correct, in that for every noun we can predict the plural form. What the description fails to express, however, is that, somehow, the three alternants [z], [s] and [iz] are the same morpheme, or, more precisely, the same morpheme alternant. It is not adequate to say that this identity is expressed in the semantics. Morphemes that mean the same thing can have
different phonological forms, in which case they are allomorphs. The English comparative is expressed by the suffix -er in some adjectives, like nice, but by means of the ‘periphrastic’ more in the case of others, like esthetic, or, again, the regular plural suffix [z] is an allomorph by the side of the plural suffix [ən] in oxen. The availability of two levels of representation makes it possible to state that underlingly the phonological form of the regular plural suffix is [z], even though in surface representations it shows up as [z], [s] and [ɪz] (Anderson 1974). Phonologically motivated morpheme alternation can thus be characterized as variation in the shape of the same underlying form in a way that differentiates such variation from cases in which the different phonological forms are unrelated, like more and -er, as well as from irregular forms, like went, the past tense form of go, for example, or feet, the plural form of foot. In such cases, the unpredictable allomorphs must be listed in the lexicon, since there is no plausible phonological generalization which could account for them.

While the case for a single underlying form for the English plural is intuitively very clear, there has been no answer to the general question when different forms should be related to a single underlying form and when they must be listed as separate word forms. Chomsky and Halle (1968) derived words like sane and sanity from the same underlying form [sæn], just as profound and profundity are derived from a common [prof ūnd] (where the overbar indicates tenseness). Others have questioned whether this is realistic, and have attempted to develop experimental procedures to answer such questions empirically (McCawley 1986; Ohala 1986; Wang and Derwing 1986).

Q64 Which of the following pairs of English words would you say contain a common underlying form?

- sew – sewage
- cork – corkage
- blow – blew
- conceive – conception
- talk – talked
- fraternal – brother

The third reason for the postulation of an underlying representation is that many generalizations are only valid at a level other than the surface level. Kenstowicz (1994a: 72), citing Mohanan (1992), gives as an example the generalization that sequences of sibilants are broken up by a vowel in English, which is responsible for the fact that the plural of English bus is [ˈbʌsɪz] rather than *[bʌsz] (or *[bass]). The generalization also holds in Singapore English: the plural of kiss is kiss[əz] and that of nose is nos[əz]. In this variety of English, there is the further fact that plosives are deleted in the coda after fricatives, so that lift, list, task are [lɪf, lɪs, tɑːs]. When these words are pluralized, they come out as [lɪfs, lɪss, tuːs], respectively. Apparently, the
generalization that adjacent sibilants must be separated by a vowel does not hold at the surface level in Singapore English, but is true at a level of representation at which the final plosive must still be present. Clearly, if we took it to be true at the surface level, the generalization would predict that the plural of list, with its surface pronunciation [lɪss], was *[ˈlɪsəz]. If instead we assume that, underlingly, the form [list] exists, it quite happily allows [z] to be added, without the need for an inserted [ə]. There is in fact independent evidence that the plosives exist at a deeper level of representation, because they show up in verbal forms before vowel-initial suffixes, as in lif[t]ing, lis[t]ing, tas[k]ing. (The assumption here is that the noun list and the verb list have the same phonological form.) In (2), the situation is schematized. The plural morpheme [z] is attached to the underlying forms of kiss and list. The [ə] is inserted between the adjacent sibilants in the form for kisses, but not in that for lists, since in the latter form the sibilants are separated by [t]. Only after the vowel has been supplied can [t] be deleted. From (2), it is clear that the generalization θ-insertion is only true for the underlying representation, not for the surface representation.

(2) Underlying:  

<table>
<thead>
<tr>
<th>θ-insertion</th>
<th>ə</th>
<th>(n.a.)</th>
<th>(n.a.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>t-deletion</td>
<td>Ø</td>
<td>(n.a.)</td>
<td></td>
</tr>
<tr>
<td>Other rules</td>
<td>(n.a.)</td>
<td>s</td>
<td>(n.a.)</td>
</tr>
<tr>
<td>Surface</td>
<td>kɪsəz</td>
<td>lɪs</td>
<td>listŋ</td>
</tr>
</tbody>
</table>

Q65

1 If it were assumed that the underlying form of the verb list was [lɪs], what incorrect prediction would be made about the pronunciation of the present participle form of this verb?

2 The past participle of regular Singapore English verbs is formed by suffixing [d] to the stem, as in [sɑːd] sɪghed, [fɑːd] fɪled. Like the plural suffix [z], it is devoiced after voiceless obstruents, as in [wɔːk] wɔːk, while [ə] is inserted if the stem ends in [t d]. If the verb list, whose past participle form is [listəd], were underlingly [lɪs], what incorrect prediction would be made about its past participle form? Why does the assumption of underlying [list] make the correct prediction?

Surface forms that contradict a phonological generalization, like [lɪss] in the last line of (2), are opaque, that is, non-transparent. Taking a rule-based perspective, opacity can arise because some rule has failed to apply and the rule’s structural description is met in the surface form. This occurs in the case of [lɪss], in which θ-insertion appears ineffective. Equally, a rule may apply even though the context of the rule is not – any longer! – present in the surface form. An example of this type
of opacity may occur in Japanese, when a vowel is deleted after causing a preceding [t] to be [ts] (cf. Q27). When the underlying form [tokemono] is pronounced [tskemono], the [u] that triggered the affrication of [t] has disappeared between voiceless consonants. Opacity is typically dealt with quite adequately in a rule-based description, because rules can be ordered so that segments can be allowed to have active or passive effects before they are deleted. In Optimality Theory, however, opacity is not easily dealt with, not, that is, if all constraints are by definition valid for output forms (see chapter 13).

7.4 NEUTRALIZATION

In the case of British English ‘dark l’, the rule we postulated produces a novel segment: before the application of the rule, no morpheme contained that segment. However, it is frequently the case that the output of a rule is a segment that already exists in the context concerned. This is true, for instance, for the rule of devoicing that devoices the [z] of the English plural suffix to [s] after voiceless obstruents, as when [bok-z] (books) is changed to [boks]. English already has such forms underlyingly, like box [bɔks]. It also holds good for final devoicing, which exists in German and Dutch, among many other languages. These rules produce segments (voiceless obstruents) in positions where such segments already appear. In a Dutch word like [pad] ‘toad’, the final [d] will be devoiced to [t], because the consonant occurs in the coda of the syllable. As a result, the singular form is pronounced [pat], but when the plural suffix [ən] is attached, the form is [ˈpadən]. Since [t] appears in that position in words like [kat] ‘cat’ (whose plural is [ˈkatən]), the opposition between [d] and [t] is neutralized in syllable-final position. Rules like English devoicing and Dutch final devoicing, therefore, are neutralization rules, while British English l-velarization is an allophonic rule.

7.5 CHOOSING THE UNDERLYING FORM

When a morpheme has a number of alternants, one of these will have to be chosen as the underlying form by the phonologist and – if the phonological model reflects our mental world – by the infant acquiring his or her language. A good underlying form satisfies two requirements. First, it should allow you to write rules that do not destroy a segmental contrast of the language. Second, the rule or rules that are needed to produce the allophony or allomorphy will be easy to state. As for the first requirement, remember that a neutralization occurs whenever a rule produces an output that already existed in the context concerned. A particularly bad choice therefore would be to choose a neutralized segment. ‘Don’t lose contrasts!’ is the important advice here. Take final devoicing, discussed earlier in the last paragraph of section 3.2 and in section 6.4.1. Recall from the previous section that the Dutch morpheme for ‘toad’ has two alternants, [pad] and [pat], occurring in the plural [ˈpadən] and the singular [pat], respectively. If we assume
that the underlying form is indeed [pad], we can derive the two alternants with the help of (3), repeated from (18) in chapter 6 in prose form.

(3) FINAL DEVOICING Obstruents are voiceless in coda position.

Notice that this rule also applies vacuously in the case of [kat] ‘cat’, which has a plural form [‘køn]. The derivation is given in (4).

(4) UR pad pa.d-ən kat ka.t-ən
   FINAL DEVOICING t (n.a.) vac. (n.a.)
   Output pat pa.dən kat ka.tən

If we were to flout the advice expressed by ‘Don’t lose contrasts!’, we might assume that there is a voicing rule that voices obstruents in the plural. Let’s assume for the sake of argument that the underlying form of ‘toad’ is [pat]. Instead of a rule devoicing obstruents in the coda, we would need a rule that voices obstruents in the onset, i.e. (5). A situation would arise which is shown in (6). Clearly, this description produces the wrong results in the case of the plural of ‘cat’. Rule (5) is not a correct generalization about Dutch, and the description in (4) must be considered superior.

(5) ONSET VOICING Obstruents in the onset are voiced.

(6) UR pat pa.t-ən kat ka.t-ən
   ONSET VOICING (n.a.) d (n.a.) d
   Output pat pa.dən kat *ka.dən

The same reasoning applies to the choice between the [s] and [z] allomorphs of the English plural suffix, discussed in section 4.4.1. With /z/ as the underlying form, we need a devoicing rule to take /bæk-z/ to [bæks]. But if we were to assume /s/ as the underlying form, a voicing rule would apply so as to change [s] to [z] after voiced segments. At first sight, this may seem fine: a form like /bok-s/ can now surface unchanged as [boks], while forms like /pen-s/ pens and /kaʊ-s/ cows are changed to [penz] and [kaʊz] by the voicing rule. However, [s] and [z] contrast after voiced sounds, as shown by dose [dəʊz] versus doze [dəʊz] and by pence [pens] versus lens [lɛnz], or indeed pens [pɛnz]. Taking the neutralized [s] as underlying would therefore lead to the incorrect voicing of [s] in pence /pɛnz/ to *[pɛnz], along with the correct creation of [penz] out of /pɛn-s/. Taking the underlying form to be /z/ means that we can safely devoice it to [s], since there is no contrast with [z] that we need to worry about.

The second requirement is simplicity in the formulation of the rules or constraints. A correct choice of the underlying form leads to correct linguistic generalizations. Thus, it is correct to say that all obstruents are voiceless in the coda in Dutch, and, conversely, it is incorrect to say that all obstruents are voiced in the onset in Dutch. Similarly, it is correct to say that all word-final coronal fricatives in English are voiceless after voiceless sounds, but it is incorrect to say that all word-final coronal fricatives are voiced after voiced sounds.\(^1\) In order to illustrate that
correct generalizations are simple to state, often simpler than if we took another morpheme alternant as the underlying form, let us consider the choice between /ɪz/ and /z/. The context for the insertion rule is simpler than that of the deletion rule. Insertion occurs ‘between sibilants’, deletion between ‘plosives, nasals, approximants, vowels and non-sibilant fricatives on the one hand and sibilants on the other’. We saw in chapter 5 how incongruous groupings are difficult to state in formal notation. The fact that contexts in rules and constraints are simple to state must reflect the fact that infants acquiring the language make generalizing hypotheses about the grammar.

The choice between /ɪz/ and /z/ could equally be motivated on the basis of the first requirement, that of no loss of contrasts, but the effect is a little more difficult to see. If we choose /z/ and an insertion rule, we correctly predict that there are no sequences of sibilants at the end of the word. That is, English has no words like *[kɪs], *[sɪʃ] or *[tɪz]. However, if we choose /ɪz/ and a rule of ɪ-deletion, we predict that there are no words ending in non-sibilants and [ɪ], like Los Angeles [lɒs ˈændʒəlz] (which exists by the side of [lɒs ˈændʒəlɪz], Wells 2008). With ɪ-deletion, we would end up with *[lɒs ˈændʒəlɪz] as one of the names of the city. Clearly, we wouldn’t want to give up the contrast between /ɪz/ and /z/ after non-sibilants.

Q66

[hwən dət aːprəl wiθ hiz fuːəz sətə]
_when that April with his showers sweet_
[ðə druːxt ɔv marʧ haθ pɜːsəd tə də rətə]
_the drought of March has pressed to the root_

are the first two lines of Geoffrey Chaucer’s _The Canterbury Tales_, with word-by-word glosses. The pronunciation of his words has been changed drastically by the many generations of speakers that have acquired the language since he wrote them some 625 years ago. For instance, the consonant [x] disappeared after making the vowel before it long, and Chaucer’s [uː] corresponds to [əʊ] today (i.e. [druːt] went from [drʊːt] to [draʊt]), and diphthongs before [r] acquired an extra syllable from the schwa-like transition between [əʊ] and [r] (i.e. [fuːr] went from [fəʊr] to [ˈfəʊr]).

1 There are two words that suggest that the present-day morpheme alternation for the plural suffix and for the past/participle suffix did not exist in Chaucer’s time. What are they?
2 Argue that the answer to the previous question shows that synchronic grammars do not necessarily reflect the historical processes of language change.
In practice, it is usually easier to see whether our choice of underlying form leads to a good generalization than whether it preserves the language’s contrasts. The place of articulation of nasal consonants in the syllable coda alternates in many languages as dictated by the place of articulation of the following consonant. In English, for instance, the preposition in is [ɪn] before the [t] of Tallinn, [ɪm] before the [p] of Paris and [ɪn] before the [k] of Copenhagen. The rule is neutralizing in the case of [m] and [n]: original [m] exists in di[m] Paris, which is di[m] in isolation, and original [n] exists in lo[n] Copenhagen, which is lo[n] in isolation. Nothing merges with [n], as there is no assimilation in di[m] Turin and lo[n] Turin. On the basis of these facts, we must avoid including a neutralizing [m] or [n] in the underlying form, and instead take [n], to give the UR /m/. But if we had incorrectly taken /m/, say, as underlying, we would probably have been alerted that something was wrong once we got to writing the assimilation to coronal [n]. This would occur before coronal [t d s z n l r], as in in Tallinn, in Dublin, in Stockholm, in Zagreb, in Nicosia, in London, in Rome, but oddly also before vowels, as in in Athens. An odd grouping like ‘coronal consonants and vowels’ cannot be properly described with the help of distinctive features. Clearly, a rule taking either /m/ or /n/ to [n] will be more complex that one taking /n/ to either [m] or [n].

Allophonic variation, too, should be described with ‘correct linguistic generalizations’. In Q60, we saw that in Tolitoli [l] appears after front vowels and at the word beginning, while [l] occurs after back vowels. It is thus simpler to state when /l/ is realized as [l] than it is to state when /l/ is realized as [l]. For this reason, /l/ is the better choice for the underlying segment. The more complex context can then be given, or implicitly understood, as ‘elsewhere’. This is shown in (7).

(7) Tolitoli I-flapping  
/l/ → [l] after back vowels  
[I] elsewhere

Q67 Both [ˈtʌn] and [ˈtom] are common Dutch first names, derived from Anthonius and Thomas, respectively. If the child’s surname was [ˈble:kvɛld] Bleekveld and the choice was between those two names, which would you advise the parents to choose for their new son?

7.6 CONCLUSION

The recognition of two levels of representation, a surface representation (SR) and a more abstract underlying representation (UR), is the cornerstone of phonological theory. It makes it possible to describe morpheme alternants as variants of the same morpheme, and opens the way to a description in which the differences in phonological form between the alternants are expressed in terms of general statements about contextually defined phonological adjustments. URs and SRs usually
differ from each other in that URs are more detailed than SRs, but SRs may also obliterate distinctions that exist in URs, and thus neutralize contrasts. There is no easy algorithm which, given a range of surface morpheme alternants, will lead to the ‘correct’ UR. URs are chosen so that the resultant grammar is the simplest that can be constructed and no incorrect predictions are made. Another question that is hard to answer is whether forms are to be regarded as morpheme alternants that have a common UR or as forms that are listed separately in the lexicon. In practice, these questions will not often vex the phonologist, since they do not seem too difficult to answer in the majority of cases. It may be expected that psycholinguistic research will provide new insights here, in particular where the second question is concerned.

Q68 The Balantak nouns in the first column of the first data set have the derived forms in the second column when prefixed with an affix meaning ‘one’. The second data set shows verbs stems and their derivations with a prefix meaning ‘unintentionally’ (Busenitz and Busenitz 1991; some of the data are inferred).

<table>
<thead>
<tr>
<th>Noun stem</th>
<th>Prefixed form</th>
<th>Gloss for stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>a wuras</td>
<td>sawuras</td>
<td>‘seed’</td>
</tr>
<tr>
<td>b bituʔon</td>
<td>sambituʔon</td>
<td>‘month’</td>
</tr>
<tr>
<td>c loloon</td>
<td>saloloon</td>
<td>‘thousand’</td>
</tr>
<tr>
<td>d taʔ</td>
<td>santaʔ</td>
<td>‘word’</td>
</tr>
<tr>
<td>e koen</td>
<td>saŋkoen</td>
<td>‘head of grain’</td>
</tr>
<tr>
<td>f utok</td>
<td>saŋutok</td>
<td>‘brain’</td>
</tr>
<tr>
<td>g sumpir</td>
<td>sansumpir</td>
<td>‘beard’</td>
</tr>
<tr>
<td>h apu</td>
<td>saŋapu</td>
<td>‘fire’</td>
</tr>
<tr>
<td>i noa</td>
<td>sanoa</td>
<td>‘breath’</td>
</tr>
<tr>
<td>j gampal</td>
<td>saŋgampal</td>
<td>‘underlayer’</td>
</tr>
<tr>
<td>k no:r</td>
<td>saŋo:r</td>
<td>‘nose’</td>
</tr>
<tr>
<td>l malom</td>
<td>samalom</td>
<td>‘night’</td>
</tr>
<tr>
<td>m roon</td>
<td>saroon</td>
<td>‘banana leaf’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb stem</th>
<th>Prefixed form</th>
<th>Gloss for stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>n giok</td>
<td>tonjgiok</td>
<td>‘move’</td>
</tr>
<tr>
<td>o pi:leʔ</td>
<td>tompi:leʔ</td>
<td>‘see’</td>
</tr>
<tr>
<td>p joːŋ</td>
<td>tojoːŋ</td>
<td>‘shake’</td>
</tr>
<tr>
<td>q kana</td>
<td>tonkana</td>
<td>‘hit’</td>
</tr>
<tr>
<td>r wawau</td>
<td>towawau</td>
<td>‘do’</td>
</tr>
<tr>
<td>s tobok</td>
<td>tontobok</td>
<td>‘stab’</td>
</tr>
<tr>
<td>t luaʔ</td>
<td>toluaʔ</td>
<td>‘vomit’</td>
</tr>
<tr>
<td>u sosop</td>
<td>tonsosop</td>
<td>‘suck’</td>
</tr>
</tbody>
</table>
Underlying and surface representations

v ƞoap  \toƞoap  ‘yawn’
w dawo?  \tɔndawo?  ‘fall’
x balo  \tɔmbalo  ‘throw’
y tunu  \tɔntunu  ‘burn’
z roƞɔr  \tɔrɔƞɔr  ‘hear’
aa u:s  \tɔnuf:s  ‘chew’

1 List the alternants of each prefix.
2 For each alternant, list the initial segment of each base before which the alternant occurs.
3 For each prefix, decide which alternant is the UR. Motivate your choice with reference to forms f, h and aa.
4 Two rules are needed to derive the surface forms. One is a place assimilation rule for nasals. The second is either a nasal insertion rule or a nasal deletion rule. Motivate your choice between these two possibilities.
5 Give a formal notation of the rule you have argued for.

Q69 The underlying and surface forms of representative Dutch noun stems and diminutive forms are listed below (e.g. van de Weijer 2002; van der Hulst 2008).

<table>
<thead>
<tr>
<th>Letter</th>
<th>Underlying Form</th>
<th>Surface Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>εi</td>
<td>εicə</td>
<td>‘egg’</td>
</tr>
<tr>
<td>b</td>
<td>lip</td>
<td>lipjə</td>
<td>‘lip’</td>
</tr>
<tr>
<td>c</td>
<td>dроyf</td>
<td>dроyfə</td>
<td>‘grape’</td>
</tr>
<tr>
<td>d</td>
<td>kra:χ</td>
<td>kra:χjə</td>
<td>‘collar’</td>
</tr>
<tr>
<td>e</td>
<td>rak</td>
<td>rakjə</td>
<td>‘skirt’</td>
</tr>
<tr>
<td>f</td>
<td>ko:нəŋ</td>
<td>ko:нəŋkjə</td>
<td>‘king’</td>
</tr>
<tr>
<td>g</td>
<td>slαŋ</td>
<td>slαŋcə</td>
<td>‘snake’</td>
</tr>
<tr>
<td>h</td>
<td>stul</td>
<td>stulcə</td>
<td>‘chair’</td>
</tr>
<tr>
<td>i</td>
<td>bοm</td>
<td>bοmcə</td>
<td>‘bomb’</td>
</tr>
<tr>
<td>j</td>
<td>kar</td>
<td>kαrcə</td>
<td>‘cart’</td>
</tr>
<tr>
<td>k</td>
<td>be:зamp</td>
<td>be:зampjə</td>
<td>‘broom’</td>
</tr>
<tr>
<td>l</td>
<td>ra:m</td>
<td>ra:mpjə</td>
<td>‘frame’</td>
</tr>
<tr>
<td>m</td>
<td>bal</td>
<td>balcə</td>
<td>‘ball’</td>
</tr>
<tr>
<td>n</td>
<td>snor</td>
<td>snorcə</td>
<td>‘moustache’</td>
</tr>
<tr>
<td>o</td>
<td>o:r</td>
<td>o:rcə</td>
<td>‘ear’</td>
</tr>
<tr>
<td>p</td>
<td>sχun</td>
<td>sχuncə</td>
<td>‘shoe’</td>
</tr>
<tr>
<td>q</td>
<td>pas</td>
<td>pafə</td>
<td>‘step’</td>
</tr>
</tbody>
</table>

1 The last rule to affect the surface forms assimilates laminal coronals (i.e. [t d s z n]) before [j], to [c ʃ ʒ ɲ], respectively. The coronal consonants [r l] are unaffected. If we take a follow-up rule of j-deletion for granted, how should palatalization be formulated?
2 List the five alternants of the diminutive suffix in their shapes before palatalization and j-deletion applied. If we take the UR of the diminutive suffix to be [tʃə], there are three rules that will be needed: t-deletion, which deletes the [t] in the suffix; ə-insertion; and place assimilation. Determine the context for t-deletion.

3 Determine the context for ə-insertion. Recall that Dutch has sets of tense and lax vowels, as shown in (6) in chapter 6.

4 Give a prose version of place assimilation. Why must this rule apply after ə-insertion?

5 Underlying forms of nouns that end in [t] end up without [t] in either the stem or the diminutive suffix, as shown in the forms below. Assume that [t] in the stem is deleted by degemination. For each of the forms s, w and x, determine a unique pair of ordered rules among degemination, t-deletion, ə-insertion and place assimilation.

| r   | naːlt | naːlcə | ‘needle’ |
| s   | mant  | mɑŋcə | ‘basket’ |
| t   | fut   | fucə  | ‘foot’   |
| u   | bɔrt  | bɔrcə | ‘plate’  |
| v   | kast  | kɑʃə | ‘cupboard’ |
| w   | bɔχt  | bɔχʃə | ‘bend’   |
| x   | hɛmt  | hɛmpʃə | ‘vest; shirt’ |

Q70 The underlying form of the diminutive suffix in the dialect of Utrecht is [ti]. Noun stems and diminutive forms are listed below (van den Berg 1975).

| a   | ei   | eiχi | ‘egg’  |
| b   | lɪp  | lɪpi | ‘lip’  |
| c   | droeyf | droeyfi | ‘grape’ |
| d   | kraːχ | kraːχi | ‘collar’ |
| e   | rɔk  | rɔki | ‘skirt’ |
| f   | koːnaŋ | koːnaŋki | ‘king’ |
| g   | slaŋ | slaŋχi | ‘snake’ |
| h   | stul | stultsi | ‘chair’ |
| i   | bɔm  | bɔmχi | ‘broom’ |
| j   | kar  | karχi | ‘cart’ |
| k   | beːzɔm | beːzɔmpi | ‘broom’ |
| l   | raːm | raːmpi | ‘frame’ |
| m   | bal  | balχi | ‘ball’ |
| n   | snɔr | snɔrχi | ‘moustache’ |
| o   | oːr  | oːrtsi | ‘ear’ |
### Underlying and Surface Representations

| p  | sχun | sχuntsi | ‘shoe’ |
| q  | pas  | pasi    | ‘step’ |
| r  | na:lt| na:ltsi | ‘needle’ |
| s  | mant | mantsi | ‘basket’ |
| t  | fut  | futsi   | ‘foot’ |
| u  | bɔrt | bɔrtsi | ‘plate’ |
| v  | kast | kasi    | ‘cupboard’ |
| w  | bɔxt | bɔxi    | ‘bend’ |
| x  | hɛmt | hɛmpi   | ‘vest; shirt’ |

1. List the alternants of the diminutive suffix.
2. The Utrecht grammar has t-deletion, degemination, ə-insertion and place assimilation, just as in Standard Dutch. Two specifically Utrecht rules are needed to account for the alternations in the diminutive suffix. Determine the position of each of these two rules in the ordered set of rules. Formulate these rules, using symbols for segments rather than distinctive features. (Hint: Their formulation can be very simple if they are ordered right.)
3. Show the derivation of items a, e, i, l, t, u and w.

### NOTE

1. In fact, the English generalization is much more general than this. It holds that no voicing differences exist within sequences of obstruents inside the word.
8.1 INTRODUCTION

In this chapter, we consider the question of how many phonological representations a word has. So far we have postulated two levels of representation, an underlying one and a surface one, a position that was motivated in chapter 7. One indication that this two-level model is inadequate is that it fails to account for the intuitions of native speakers about the pronunciation of the words of their language. In brief, the underlying representation would seem to be too abstract, while the surface representation appears to be too detailed (cf. Schane 1971). The advent of Lexical Phonology put an end to this particular shortcoming. It postulates an intermediate level of representation, the lexical representation (Kiparsky 1982a, 1985, 1993; Mohanan 1986). Not only does it correspond to native-speaker intuitions, it will also be shown to have a number of interesting properties. After dealing with this three-level phonological model, we turn our attention to the relation between the surface representation and the physical pronunciation of the words. We will see that languages typically differ in the way they realize surface representations which must be assumed to be identical. Such language-specific realizations of phonological elements are accounted for by rules of phonetic implementation.

8.2 DEFINING AN INTERMEDIATE LEVEL OF REPRESENTATION

If you look up the pronunciation of a word in a dictionary, you will find it is normally given in phonemic transcription. In this type of transcription, an English word like *pin* is transcribed [pɪn], not [pʰɪn]. This is because the segment [pʰ] is an allophonic variant of the phoneme /p/, and as such has no place in a phonemic transcription. Clearly, the dictionary’s phonemic transcription defines a level of representation which is more abstract than the surface level. However, it does not correspond with the underlying representation either. The pronunciation of *looked*, for instance, would be given as [lukt], not as [lokd]. The form that is given incorporates the output of a devoicing rule that makes obstruents voiceless after voiceless obstruents. That is, the phonemic transcription apparently corresponds to a level of representation which is somewhere between the underlying level and the surface level.

It might at first sight seem reasonable to suppose that the intermediate level corresponds to the output of all rules that produce existing segments, which were called ‘neutralization rules’ in chapter 7, while all rules that produce novel segments, or
‘allophonic rules’, would then apply to the intermediate representation so as to produce the more detailed surface representation. This would put all the rules that produce phonemes, or existing segments, in a different compartment of the grammar from all the rules that produce allophones, or novel segments. While this assumption is almost correct, there is still something not quite right. This is because there are rules whose output is a mix of existing and novel segments. Not all rules allow themselves to be characterized as either ‘neutralizing’ or ‘allophonic’. Some rules are both: depending on the input, they produce either an already existing segment or a novel segment. For instance, Dutch has a rule that voices plosives before [b d], called regressive voicing. (It applies to all obstruents, but we will ignore the fricatives here.) The language has the three voiceless plosives [p t k], each of which can appear before /b d/. REGRESSIVE VOICING thus produces the voiced plosives [b d g]. Dutch has oppositions between /p/ and /b/, as in [pak] ‘parcel’ – [buk] ‘tray’, and /t/ and /d/, as in [tak] ‘branch’ – [dak] ‘roof’, but there is no contrast /k/–/g/. Therefore, the output of regressive voicing is partly phonemic, viz. when /p t/ are voiced to /b d/, and partly allophonic, viz. when /k/ is voiced to /g/. In (1), the first column gives the underlying representations, while the third column gives the results of regressive voicing.

<table>
<thead>
<tr>
<th>(1) Underlying</th>
<th>Phonemic</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>ap dun</td>
<td>ob dun</td>
<td>[ob dun] ‘put on’</td>
</tr>
<tr>
<td>oeyt bra:k</td>
<td>oeyd bra:k</td>
<td>[oeyd bra:k] ‘breakout’</td>
</tr>
<tr>
<td>zakt buk</td>
<td>zak buk</td>
<td>[zak buk] ‘pocket book’</td>
</tr>
</tbody>
</table>

It should be clear that it would not be very clever to maintain a distinction between rules that produce existing segments and rules that produce novel segments, for this would mean that we would have to split regressive voicing, given in (2), into two rules, one to produce the existing segments /b d/, the ‘neutralizing’ part of the rule, shown in (3a), and a second to produce the novel [g], the ‘allophonic’ part, shown in (3b). (We formulate the rules so as to voice all obstruents.)

\[
\begin{align*}
\text{(2) Regressive voicing} & : [-\text{son}] \rightarrow [+\text{voice}]/__\text{(##)} [\begin{array}{l}
-\text{son} \\
-\text{cont} \\
+\text{voice}
\end{array}] \\
\text{(3) a Regr. p, t-voicing} & : [-\text{son}] \rightarrow [+\text{voice}]/__\text{(##)} [\begin{array}{l}
-\text{son} \\
-\text{cont} \\
+\text{voice}
\end{array}] \\
\text{(3) b Regr. k-voicing} & : [-\text{son}] \rightarrow [+\text{voice}]/__\text{(##)} [\begin{array}{l}
-\text{son} \\
-\text{cont} \\
+\text{voice}
\end{array}]
\end{align*}
\]

Rule (2) converts the underlying forms in the first column in (1) into the forms in the third column, which seems just right. If we wanted to make a clear cut between a phonemic module and an allophonic module in the phonological grammar, rule (2)
would have to be split up into (3a), the neutralizing, phonemic rule (to produce /b d/) and (3b), the allophonic rule (to produce [g]).

Halle (1959) used the above argument, which he made on the basis of a similar case of undesirable rule duplication in Russian, to argue that the notion of the ‘Structuralist Phoneme’, which was a cornerstone of phonological theory as it existed before the advent of ‘Generative Phonology’, was misguided. The twist lies in whether we think of [d] as produced by (2) as being ‘a pronunciation of /d/’ or ‘a pronunciation of /t/’. American Structuralist phonologists like Zellig S. Harris, Charles F. Hockett and W. Freeman Twaddell, whose work was superseded by that of the Generativists of the late 1950s and 1960s, like Morris Halle, Noam Chomsky and Paul Postal, held that the underlying representation was transferred into an intermediate, phonemic representation, which in its turn was transformed by rules into an allophonic surface representation. The notion phoneme is based on the existence of a surface contrast. The (entirely reasonable and valid) idea was that whenever a minimal pair could be formed, i.e. a pair of words whose members differ by one segment only, like English lock – rock, or beat – boot, the two segments responsible for the difference must belong to different phonemes. So the above two minimal pairs are evidence for the existence of the phoneme categories /l/ – /t/ and /i:/ – /u:/ in English. What made their theory problematic, however, was the additional premiss that a segment could only belong to one phoneme. So once a /t/ had been set up in a language on the basis of some minimal pair in which the segment [t] contrasts with some other segment, any occurrence of [t] in any other word, regardless of context, also belonged to the phoneme /t/. This premiss has been referred to as ‘Once a phoneme, always a phoneme’. Adherence to this principle leads inevitably to the sort of undesirable splits in generalizations that we have seen in the case of Dutch regressive voicing. Since [d] belongs to the phoneme /d/, a conclusion based on minimal pairs in which it contrasts with [t] in the onset, [d] must also be assumed to represent /d/ when it occurs in the coda before a voiced plosive, as in the word for ‘breakout’ in (1). Since the same reasoning does not apply to [g], the position encapsulated in ‘Once a phoneme, always a phoneme’ implies the recognition of an undesirable representational level illustrated in the second column of (1).

Halle’s argument makes it clear that there is no place for the Structuralist Phoneme in phonological theory. It does not, of course, rule out the existence of any intermediate representation. The answer to the question of what the appropriate intermediate level is has been provided by the theory of Lexical Phonology.

8.3 LEXICAL PHONOLOGY

The crucial assumption made by Lexical Phonology is that some of the phonological generalizations of a language are stated in the lexicon, the morphological module which incorporates the semantic, phonological and morphological information of the language’s morphemes, while others are stated outside it. That is, a distinction is drawn between lexical phonological rules and postlexical phonological rules,
the latter applying after the words have been inserted into the phrase (Kiparsky 1982a; Mohanan 1986). In (4), a number of the distinguishing properties are listed.

(4) | **Lexical rules** | **Postlexical rules** |
---|---|---|
| a. May refer to morphological labels | Cannot refer to morphological labels |
| b. May have exceptions | Cannot have exceptions |
| c. Structure-preserving | Need not be structure-preserving |
| d. Accessible to native-speaker intuition | Not easily accessible to native-speaker intuition |
| e. Cannot apply across word boundaries | May apply across word boundaries |
| f. Must precede all postlexical rules | Must follow all lexical rules |

We discuss each of these properties in the following sections.

### 8.3.1 Reference to morphological labels

Since lexical rules apply inside the lexicon and postlexical rules do not, only the former have access to category labels like ‘N(oun)’, ‘V(erb)’, etc. For example, there is a rule in Dutch that deletes word-final [n] after [ə], as shown in (5).

(5) | **Underlying** | **Surface** |
---|---|---|
| [[loːp] + ən] | loːpə | ‘to walk’ |
| [[zaːk] + ən] | zaːkə | ‘things’ |
| [teːkən] | teːkə | ‘sign’ |
| [[teːkən] + ən] | teːkənə | ‘to draw’ |
| [oːpən] | oːpə | ‘open’ |

When [ən] occurs finally in a verb stem, however, no deletion takes place (Koe-foed 1979; Trommelen and Zonneveld 1983). This is shown in (6).

(6) | **Underlying** | **Surface** |
---|---|---|
| [teːkən] | teːkən | ‘draw’ |
| [oːpən] | oːpən | ‘open’ |

**n-deletion** (7) thus distinguishes between the noun ‘sign’ and adjective ‘open’ in (5) on the one hand and the verb stems for ‘draw’ and ‘open’ in (6) on the other. Such a condition can only be put on a lexical rule. Once a word has left the lexicon and has been inserted in syntactic structure, category labels are removed, and postlexical rules therefore cannot refer to them.

(7) **n-deletion**

\[ n \rightarrow \emptyset / \_ \_ \_ \_ \_ \]  

Condition: \( X \neq \text{Verb} \)

### 8.3.2 Exceptions

Lexical rules, but not postlexical rules, have access to the lexicon, and as such can tell which word they are dealing with. A rule that has exceptions, therefore, cannot be a postlexical rule. To return to the example of **n-deletion** (7) above: [ˈheɪdən]
'heathen' and ['krıːstən] 'Christian' are exceptional in not undergoing the rule. The entries of these words are assumed to be provided with the information 'Not subject to (7)'. Similarly, English has a rule of trisyllabic laxing, which laxes a vowel in the antepenultimate syllable of words derived with suffixes like -ity. Examples are given in (8). However, the words nicety and obesity (cf. nice, obese) exceptionally have [æ iː:] in the antepenultimate syllable, rather than the expected lax [i ɛ].

\[
\begin{array}{ccc}
\text{Tense} & \text{Lax} \\
div[æ]ne & div[ɛ]nty \\
v[ɛ]n & v[æ]nty \\
ser[iː]ne & ser[ɛ]nty \\
\end{array}
\]

While postlexical rules cannot have exceptions, lexical rules could either have exceptions or be exceptionless. For instance, English has a rule deleting [n] after [m] at the end of the word (Kiparsky 1985). This rule must be lexical, because it needs information about the status of the word before inflectional endings are added. Thus, it applies in the first column in (9), and in the second column, where the words have been provided with inflectional endings, but not in the third column, where [n] is not final in the word. This lexical n-deletion rule is exceptionless: there are no words in English that end in [mn].

\[
\begin{array}{ccc}
\text{Stem} & \text{Inflected form} & \text{Derived form} \\
\end{array}
\]

8.3.3 Structure preservation

Lexical rules are structure-preserving in the sense that their output is confined to segments that already exist in underlying representations. The idea is that there is a lexical inventory of vowels, consonants and tones which is smaller than the inventory observable in surface representation. For example, since in the underlying representation of English words there is no need to distinguish aspirated from unaspirated plosives, this distinction being allophonic in English, the rule that creates aspirated plosives must be postlexical. The segments [pʰ tʰ kʰ] are novel segments, i.e. not included in the English lexical segment inventory.

Structure preservation is not an exceptionless property of lexical rules. A number of varieties of English have rules that apply before the affixation of inflectional endings and must for that reason be lexical, like the rule that deletes [n] in words like autumn, discussed above. The point is that many such rules produce novel segments (Harris 1994: 28). An example is the Scottish rule lengthening word-final vowels (as well as vowels before [+voice, +cont] segments) (Aitken’s Law). Unexpectedly, the inflectional suffix [d] of the past tense and past participle is ignored in the context of the rule, which makes the rule a lexical rule. This is shown in (10). But the product of the rule is a novel segment.
Interestingly, there is evidence that when a novel sound is produced by a lexical rule, it may be made available for inclusion in lexical representations. Thus, some speakers pronounce *concise* and *scythe* with [æ], even though in Scottish English these words end in [s] and [θ], respectively (Aitken 1984). Clearly, developmental stages in which all occurrences of a novel segment are produced by a lexical rule must be expected to occur, if it is assumed that at least some lexical rules historically start out as postlexical rules.

### 8.3.4 Native-speaker intuitions

Native speakers would appear to make reference to the lexical representation when determining whether two phonetically different sounds are ‘the same sound’; their judgements refer to the lexical segment inventory. For instance, native speakers of English regard the second segment in *stop* and the first segment of *top* as the same sound, even though they are phonetically different, which fits with the assumption that aspiration is a postlexical rule. Likewise, phonetically identical sounds that were neutralized by a postlexical rule will typically be looked upon as different sounds. As a result of the American English process of *flapping*, the intervocalic consonants in *Adam* and *atom* are phonetically identical in all styles except the most formal ones ([ˈærəm]), but native speakers nevertheless consider them different consonants. By contrast, when a lexical rule neutralizes an underlying opposition, the intuition of the native speaker tends to conform to its output.

### 8.3.5 Application across word boundaries

Because lexical rules apply in the lexicon, their structural description can never be determined by elements taken from different words. A rule that applies across word boundaries, therefore, must be a postlexical rule. Dutch *regressive voicing* (2) is a postlexical rule for this reason.

### Q71

1. What do you think is the lexical representation of ‘breakout’ in (1)?
2. Do you think that Dutch *regressive voicing* has exceptions?
3. What would native speakers of Dutch say is the last consonant of the prefix in [œyd-braːk]?
8.3.6 Lexical rules apply before postlexical rules

The final distinguishing property listed in (4) is once more a necessary consequence of the Lexical Phonology model. Words get inserted into postlexical structures in their lexical representations, i.e. after all lexical rules have applied. It follows that if we know that a rule is postlexical, a rule that must apply after it must also be postlexical, and thus display the postlexical properties listed in (4).

8.4 PHONOLOGICAL INFORMATION IN THE LEXICON

All phonological rules may refer to phonological information. The phonological information available in the lexicon is not confined to segments. Syllable and foot structure also exist in the lexicon of Dutch and English (Booij 1988; Inkelas 1989). Evidence for this position is provided by morphological processes that are sensitive to the syllable structure or the stress of the base. For example, the English comparative and superlative suffixes [[ ]_{Adj} ər] and [[ ]_{Adj} əst] require that the base should not exceed a binary foot. Therefore, the formation is allowed with the adjectives in (11a) but not with those in (11b).

(11) a white \( (\text{wart})_r \) \( (\text{wart}ə)r \)
   noble \( (\text{nəobl})_r \) \( (\text{nəoblər})_r \)
   silly \( (\text{sili})_r \) \( (\text{siliər})_r \)
   b beautiful \( (\text{bjuːtəfl})_r \) \( * (\text{bjuːtəflər})_r \)
   manifest \( (\text{mænə})_r \) \( *(\text{mænər})_r \)
   serene \( *\text{so(ri:n)}_r \) \( *\text{so(ri:nər)}_r \)

A possible illustration of the point that lexical rules may refer to both morphological structure and phonological structure is provided by final devoicing in two varieties of German. In Low German, the varieties of German spoken in the northern half of the country, this process applies to syllable-final obstruents, as shown in the second column of (12) (Venneman 1972). The requirement that the obstruent should occur in the coda also holds for High German, but this variety requires in addition that in positions before sonorants the obstruent should be morpheme-final. As a result, the examples in (12b), in which the obstruent is both syllable- and stem-final, show devoicing in both varieties, but differences appear in the case of such words as Adler ‘eagle’, where an obstruent is syllable-final, but not stem-final, as shown in (12c).

(12) Underlying Low German High German
a kind \( \text{kɪnd} \) \( \text{kɪnt} \) \( \text{kɪnt 'child'} \)
   ‘kind-ɪʃ’ \( \text{ˈkɪnd-ɪʃ} \) \( \text{ˈkɪn.dɪʃ} \) \( \text{ˈkɪn.dɪʃ 'childish'} \)
   ‘taːg-lɪç ˈteːk-lɪç ˈteːk-lɪç 'daily' \)
   ‘vaːgnɐʀ ˈvaːk.nɐʀ ˈvaːgnɐʀ 'Wagner' \)
   ‘magmaː ˈmak.maː ˈmagmaː 'magma' \)
   ‘aːdlɐʀ ˈaːt.lɐʀ ˈaːd.lɐʀ 'eagle' \)
   ‘ɔrdn-ʊŋ ˈɔrt.nʊŋ ˈɔrd.nʊŋ 'order' \)
   ‘kɪnt-kɪnt kɪnt kɪnt 'child' \)
   ‘kɪnd-ɪʃ ˈkɪn.dɪʃ ˈkɪn.dɪʃ 'childish' \)
   ‘teːk-lɪç ˈteːk-lɪç 'daily' \)
   ‘vaːg.nɐʀ ˈvaːg.nɐʀ 'Wagner' \)
   ‘magmaː ˈmagmaː 'magma' \)
   ‘aːd.lɐʀ ˈaːd.lɐʀ 'eagle' \)
   ‘ord.nʊŋ ˈord.nʊŋ 'order' \)
   ‘kɪnt-kɪnt kɪnt kɪnt 'child' \)
   ‘kɪnd-ɪʃ ˈkɪn.dɪʃ ˈkɪn.dɪʃ 'childish' \)
   ‘teːk-lɪç ˈteːk-lɪç 'daily' \)
   ‘vaːg.nɐʀ ˈvaːg.nɐʀ 'Wagner' \)
   ‘magmaː ˈmagmaː 'magma' \)
   ‘aːd.lɐʀ ˈaːd.lɐʀ 'eagle' \)
   ‘ord.nʊŋ ˈord.nʊŋ 'order' \)

These data have been given different accounts, though. Rubach (1990) provides an account based on resyllabification of the High German consonant sequences ([e.g. ˈvaːːgnər]). Similarly, Giegerich (1992) assumes that the plosive moves to syllable-initial position in some speech styles. A drawback to these solutions is that they run counter to the generalization that clusters which are ill-formed word-internally, like French [ps-] in capsule [kap.syl], *[ka.psyl], may be well-formed word-initially, as in psychologie (Selkirk 1982), but that the opposite case (e.g. *[dl] word-initially, but admissible word-medially in a case like Adler) has not otherwise been attested.

Q72 The Korean lexical consonant inventory is given in (1).

(1) LAB COR [+ant] COR [-ant] DORSAL LAR
b d dʒ g
p t tʃ k
pʰ tʰ tʃʰ kʰ
s
z
m n η
l

Before [i], [d t tʰ s z n l] are prepalatal by a rule of palatalization: [j c cʰ e z n ʎ], as shown in (2), causing surface syllables like *[si] to be ill-formed.

(2) gaps-i gapci ‘price+nominative’
sigan eigan ‘time’
ki ki ‘meal’
batʰi bαtʰi ‘endure’
tʰi cʰi ‘dust’
gaksi gakxi ‘bride’
madi mαdi ‘knot’

A rule of affrication causes coronal plosives to become affricates before [i], if [i] forms part of a suffix. Thus, [d tʰ] are replaced with [dʒ tʃ] in the contexts shown in (3a). ([t] happens not to occur stem-finally.) The affricates [dʒ tʃ] also appear in underlying representation, as shown in (3b) (after Kiparsky 1993; symbols as in Lee 1993).

(3) a batʰ-i batʃʰi ‘field+nominative’
gud-i gudʒi ‘harden+ADV’
gjatʰ-i gjatʃʰi ‘side+nominative’
b dʒip dʒip ‘house’
dʒidʒa dʒidʒa ‘tear (IMP)’
dʒad-i dʒadʒi ‘milk+nominative’
1. Is palatalization a structure-preserving rule?
2. Is affrication a structure-preserving rule?
3. Mention three properties of affrication that are consistent with its status as a lexical rule.

**Q73** French has a rule of vowel nasalization, which nasalizes a vowel before [n] in the same syllable (Tranel 1981). After vowel nasalization, a rule of n-deletion deletes [n] if it appears in the coda after a nasalized vowel. This is illustrated in (1), where the feminine forms were created by suffixing the stem with [ə]. (The nasal vowel [ɛ̆] is the reflex of underlying [i ñ e]; please ignore the variation between [e] and [ɛ].)

<table>
<thead>
<tr>
<th>Stem</th>
<th>Masc.</th>
<th>Fem.</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>plén</td>
<td>‘full’</td>
<td>plé</td>
<td>pleníə</td>
</tr>
<tr>
<td>bryn</td>
<td>‘brown’</td>
<td>brɛ</td>
<td>bryñə</td>
</tr>
<tr>
<td>fin</td>
<td>‘fine’</td>
<td>fɛ</td>
<td>finə</td>
</tr>
<tr>
<td>bɔn</td>
<td>‘good’</td>
<td>bɔ̃</td>
<td>bɔnə</td>
</tr>
<tr>
<td>roman</td>
<td>‘Romance’</td>
<td>ōm</td>
<td>romanɔ</td>
</tr>
</tbody>
</table>

There is a third rule, ə-deletion, which deletes foot-final [ə] except in the most formal styles, causing the feminine forms in (1) to be [plén, bryn, fin, bɔn, roman] on the surface.

1. Give formal notations of the three rules, referring to syllable and foot boundaries as in example (18) in section 6.4.1.
2. Show the derivation of the masculine and feminine forms for ‘fine’.
3. Before a consonant-initial noun, the feminine indefinite article is [ yn], as in [yn fam] ‘a woman’; before a vowel-initial feminine word like [e.ro. in] ‘heroine’, the indefinite article [yn] is divided over two syllables, as in [ye.ro.in] ‘a heroine’. The masculine form of the indefinite article is [ɛ̆] before consonant-initial nouns, as in [ɛ̆ gars] and [ɛ̆n] before vowel-initial nouns, where the [n] is syllabified with the noun-initial vowel, as in [ɛ̆nɔ̃m] ‘a man’. How would you account for the preservation of the [n] in the masculine indefinite article?
4. Why must n-deletion be postlexical?
5. In some contexts, word-final [ə] is preserved, as in [bãdə desine] ‘strip cartoon’. Is ə-deletion a lexical or a postlexical rule?
6. Citroën is pronounced [si.trœ.œn]. Why can this fact be used as an argument for the assumption that vowel nasalization is a lexical rule?
7. There are many words like [ɔðə] ‘wave’, which only ever have a nasalized vowel, never an oral vowel plus [n]. What would you assume as the underlying form of the word for ‘wave’?
8.5 CONTROVERSIAL PROPERTIES OF LEXICAL RULES

Other properties that have been claimed to distinguish lexical from postlexical rules are not as convincing as the properties mentioned in (4), and have therefore been disputed (Halle and Kenstowicz 1991; Kiparsky 1993). One of these is **non-derived environment blocking**. This is the phenomenon that many lexical rules would appear to skip underived words, i.e. to apply only to forms that are derived. For instance, English **trisyllabic laxing** applies to derived forms like [ˈvænæti] vanity, [dəˈvɪnæti] divinity, but never to underived forms like *ivory* [ˈɪvəri], *nightingale* [ˈnætɪŋɡɛl] *[ˈnɪtɪŋɡɛl]. Accordingly, it has been proposed that lexical rules can only apply to derived forms. However, many cases have been presented in which structure-preserving rules apply to all occurrences of a morpheme, derived or underived.

| Q74 In Dutch, word-final [n] is deleted after [ə] (section 8.3.1). Does the rule show the effect of non-derived environment blocking? |
| Q75 In Northern Irish English, all occurrences of [ɛː] can be derived from [iə] (Harris 1994). Some alternations are shown below. |

<table>
<thead>
<tr>
<th>[iə]</th>
<th>[ɛː]</th>
<th>[iə]</th>
<th>[ɛː]</th>
</tr>
</thead>
<tbody>
<tr>
<td>fate</td>
<td>day</td>
<td>station</td>
<td>pay them</td>
</tr>
<tr>
<td>made</td>
<td>stayed</td>
<td>same</td>
<td>say more</td>
</tr>
<tr>
<td>raise</td>
<td>rays</td>
<td>fail</td>
<td>daily</td>
</tr>
<tr>
<td>baby</td>
<td>playful</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Is the rule that produces [ɛː] a lexical rule?
2. Is it a neutralization rule?
3. Is it structure-preserving?
4. Does it show the effect of non-derived environment blocking?

8.6 BEYOND THE SURFACE REPRESENTATION

We have seen above that the phonological grammar advocated by Lexical Phonology can be schematized as in (13).
The surface phonological representation which is produced by the grammar contains all the linguistic information the articulators require to do the work of pronouncing the expressions concerned. In other words, it is a fully specified phonological representation. Converting that cognitive representation into the physiological actions that constitute the articulation of the expression is the task of the **phonetic implementation rules**. The diagram in (13) can therefore be completed as in (14).

### 8.6.1 Phonetic implementation

The phonological surface representation which ultimately results from the phonological grammar consists of some configuration of phonological features and structures. The translation of this discrete representation into quantitative physiological activity, the articulatory actions discussed in chapter 2, is taken care of by rules of **phonetic implementation** (Pierrehumbert 1980, 1990; Keating 1990). These take phonological presentations as input and deliver continuous acoustic signals as output. It has been claimed that earlier work, like Chomsky and Halle (1968), assumed that such rules were universal, but we have not found that assumption expressed in print. In any event, everyday experience tells us that this would be an improbable position. The same phonological syllable is always pronounced somewhat differently by speakers of different languages. For instance, a syllable like [kis] will typically be pronounced with a somewhat more fronted [k] in French than in either Spanish or Dutch, while [s] will have lower-frequency friction in Spanish than in French and perhaps Dutch. Phonetic implementation is not to be thought of as giving
Levels of representation

speaker-specific invariant outputs. It is not an ‘automatic’ algorithm, but a complex
set of articulatory procedures whose execution is under the control of the speaker
(Kingston and Diehl 1994). Speakers will make adjustments so as to preserve the
distinctiveness of phonological contrasts, depending on ambient noise or distance
to the hearer. In addition, a host of non-linguistic factors will affect the articulation.
 Speakers of the same language will fine-tune the articulation of certain segments in
order to signal the social group they belong to. As a result, we can often form a fair
idea of the speaker’s social background on the basis of the precise quality of their
vowels, say. Similarly, the speaker’s emotion will affect pronunciation in complex
ways, while communicative urgency will determine the precision with which speech
is produced.

8.6.2 Models of implementation

In a widely adopted view, phonetic implementation involves the translation of
a phonological feature into a target for the designated articulatory parameter.
For instance, the feature [-cont] triggers the formation of an oral closure at the
location specified by the place features. A feature like [+nasal] will trigger an
opening of the velum for the segment specified for that feature, while [-nasal]
will do the opposite. The different targets of the same articulatory parameter will
be connected up by transitions known as interpolations. Thus, the movement

Q76 In New York City English, BATH-RAISING raises and diphthongizes
[æ] to [ɛ] or [ɪ]. The rule applies in a number of phonological con-
texts, among which is ‘when a nasal follows in the same syllable’. It
thus fails in family [fæ.mə.li] but applies in damn [dəm]. Moreover,
it applies in underived major-class stems, so that the stressed minor-
class auxiliary can has [æ], while the noun can (of beer) has [ɪ]. Like-
wise, swam (a derived past tense of swim) has [æ], while ham has [ɪ].

1 Is BATH-RAISING a rule of phonetic implementation, a postlexical pho-
nological rule or a lexical phonological rule? Motivate your answer.
2 English swan [swən] used to have an unrounded [a] in the oldest stages of
the language. The [w] caused [a] to have lip rounding, after which it was
interpreted as [ɔ], unless a velar consonant followed. This also took place
in words like quality, which were borrowed from Norman French with
an unrounded [a] after the invasion of 1066. However, the [a] in present-
day words with [wa] doesn’t acquire this rounding, as shown by aquatic
[aˈkwætɪk], and there is thus no productive rule of [a]-rounding today.
How could you show that BATH-RAISING is productive? Hint: Think of a
New York school child learning a foreign language.
of the velum from a raised to a lowered position constitutes the interpolation between two phonologically specified target positions. This view of phonetic implementation implies two things. First, phonetic implementation rules do not themselves add or remove phonological features: the only thing they can do is translate features into articulations. Second, segments that have no specification for a particular feature may be in the path of an interpolation between two segments that do. For instance, Cohn (1990) shows that the nasalization of the vowel in American English words like den, room, sung should be described in just that way. The nasalization starts early in the vowel and progressively increases to reach full velic opening at the end of the vowel. This is shown graphically in (15a), where the graph represents the nasal airflow, or – indirectly – the degree to which the velum is lowered. Cohn argues that in (15a) the vowel has no specification for the feature [±nasal], and that its nasality results from the interpolation between the [–nasal] specification of the preceding consonant and the [+nasal] specification of the following consonant. By contrast, in words like camp, dent, sink, the velum opens fully right at the start of the vowel and closes fairly suddenly at the moment when the oral contact for the voiceless plosive is made. In fact, the nasal consonant disappears before a voiceless plosive in the coda, so that the nasality of the vowel can be seen as the transfer of that feature from the deleted [m n ŋ]. This is shown in (15b). (The voiceless plosive is accompanied by a glottal closure introduced by an independent rule of glottalization.) In (15c), the nasality ‘cline’ runs in the opposite direction, in words like mood, Ned. Thus, only in (15b) does the nasality of the vowel result from a [+nasal] feature on the vowel, which is supplied by a phonological rule nasalizing vowels before a sequence of a tautosyllabic (i.e. occurring in the same syllable) nasal and plosive. In (15a,c), the nasality on the vowel results from an interpolation between opposite values for [±nasal] on either side of the vowel.

An alternative view holds that phonological features are translated into gestures. That is, instead of defining the beginning and end points of articulatory movements, the elements in a gestural model are the articulatory movements themselves (Zsiga 1997). Under this view, the phonological features in the syllable [den] are translated into a laryngeal gesture (voicing), two tongue tip gestures (one for [d] and one for [n]), a tongue body gesture (palatal wide, to produce the quality of [ɛ]), and a velic closing and a velic opening gesture. Their timings are governed first by the order of the segments and second by the more detailed, language-specific instructions. If the opening gesture of the velum is initiated at the same time as the tongue tip closing gesture, there will be some nasalization during the last part of the vowel: the velum opens immediately when the gesture is started, while it will take the tongue tip some time actually to reach the alveolar ridge and, with the rims of the tongue,
make the alveolar closure. In a language that allows a greater degree of nasalization, the velic opening gesture is extended so that it starts earlier.

The developers of the gestural model had originally intended the gestures to be the phonological features encoding the pronunciation of words in the lexicon (Browman and Goldstein 1989). A problem with this ‘phonetic’ view of phonological representations is that it makes it difficult to state the contrasts a language employs. It would, for instance, be impossible to say whether the nasalization during a vowel was contrastive or allophonic. A related assumption was that phonological adjustments arose from extending the duration of gestures, such that gestures come to overlap (as in the hypothetical case of the nasalization of prenasal vowels noted above). This, however, makes it impossible to distinguish between phonological, i.e. categorical, effects (which typically cause gestures to disappear or be reordered) and gradient phonetic effects (which can be very adequately expressed by adjusting the durations of gestures). The gestural model is thus better suited for dealing with the phonetic implementation than for encoding underlying forms.

Recently, phonetic implementation has been thought of less as the construction of an acoustic signal on the basis of gesture-based or target-based interpretations of phonological features than as the probabilistic selection of auditory representations remembered by the speaker, known as ‘exemplars’. Speakers appear to modify the pronunciation of specific words on the basis of the pronunciation of other speakers whom they recently heard pronounce those words. These word-by-word adjustments suggest that we store detailed phonetic information about individual words. A well-known argument for the fact that we store minimal information about a word’s phonology is that we cannot usually say what any word of our language, piano, say, sounded like when we first heard it, and are typically unable to say whether the speaker was male or female, adult or child, or what its pitch pattern or speech tempo was (Halle 1985). Adherents of exemplar models will no doubt grant that this is true, but point out that we do remember the last time or couple of times we heard it. Tacitly or overtly, we remember a great deal about speakers and the way they said things, and we may well use that information during our own speech production. There is no conflict between the two facts, however, since one refers to our grammar, our knowledge of contrasts and the structures they are embedded in, while the other refers to how we pronounced them (Pierrehumbert 2002). Models of phonetic implementation that are based on word-based phonetic knowledge are known as ‘exemplar models’.

8.6.3 Deciding between phonology and phonetic implementation

It is not always obvious whether a particular generalization should be accounted for in the phonology or in the phonetic implementation. A reasonable assumption would be that if the output of the rule crucially feeds into another rule, the regularity must be phonological. An example may clarify this point. In many varieties of British English, including Received Pronunciation (RP), there is a tendency to insert a voiceless plosive between a nasal and a following fricative, if they belong to...
the same syllable. As a result, a word like *sense* is pronounced as if it was the plural of *cent*, i.e. as [sents] rather than [sens]. At first sight, there would appear to be two possible ways of accounting for this process. First, it might result from the details of phonetic implementation. For [n], there is an alveolar closure and the velum is lowered. To get to [s], the alveolar closure must be gradually released while the velum must be raised. The occurrence of the [t] will now depend on the precise timing of these two articulatory gestures: if the velum is raised before the alveolar closure is released, an oral stop will come to exist between the moment the velic closure is made and the alveolar closure is released. (The vocal folds will typically stop vibrating once the velic closure is made, and the stop therefore is voiceless.) This account is in principle quite plausible, but would not appear to be right in the case of RP. As it happens, the insertion of [t] creates the context for a phonological process which glottalizes voiceless stops in the coda of the syllable. If the insertion of the voiceless stop in *sense* is recorded in the phonological representation, we should expect it to trigger preglottalization. The fact that such inserted stops do trigger preglottalization, just as do underlying stops, implies that voiceless stop insertion must be a phonological process.

Q77 Would you expect *sense* and *cents* to be homophonous in British English?

By contrast, the equivalent process in American English is an instance of phonetic implementation (Ohala 1986). The first syllable in a word like *teamster* or *sensitive* may well be followed by an intrusive voiceless stop, but it does not feed into the phonology of the language: these syllables appear to be longer than syllables closed by a voiceless stop would be in the same context. Also, words like *sense* and *cents* are distinct. *Cents* has a categorically nasalized vowel, as well as a glottal stop ([sɛʔts] (see (15b)), while *sense* has a partially nasal vowel and no glottal stop, a pronunciation corresponding to (15a): [sɛn(t)s]). Thus, in an SPE-type description, the British English situation is captured by (16), while the American English situation is best described by some such instruction as in (17).

(16) **BrE stop insertion** \( \emptyset \rightarrow \begin{array}{c} \text{[–cont]} \\ \text{[–voice]} \\ \text{[αPLACE]} \end{array} / \sigma [\ldots [\text{+nas}] [\text{αPLACE}] [\text{–son}] [\text{+cont}] \ldots ] \sigma \)

(17) **AmE stop insertion** (Implementation): When realizing an intraword sequence \([+nas][-son, +\text{cont}]\) before an unstressed syllable, time the velic closure for [s] just before or at the moment of the release of the alveolar closure for [n].
Q78 Before voiced obstruents, vowels are longer than before voiceless obstruents. Languages vary in the amount of lengthening before voiced obstruents. The difference in French ([vid] ‘empty’ vs [vit] ‘quickly’) is much smaller than in English ([biːd] bead vs [biːt] beat), for instance. How would you account for this lengthening before voiced obstruents?

Q79 English [s] before [ʃ], as in stocks shelves, trace shapes, often assimilates to a consonant that is indistinguishable from [ʃ], but is also frequently pronounced as a fricative that gradually moves from [s] to [ʃ] (Holst and Nolan 1995). How would you characterize these two situations in terms of the rule typology discussed in this chapter?

8.7 CONCLUSION

We have shown that Lexical Phonology provides a satisfactory answer to the problem of the definition of a level of representation between the underlying representation and the surface representation. Instead of drawing the dividing line between rules of neutralization and rules of allophony, it assumes that phonology exists in two separate components in the grammar. One part resides in the lexicon, where it can refer to information present in the lexicon, which is, first, morphological information and, second, the set of lexical phonological segments and structures. The other part is outside the lexicon, where morphological information is no longer available and segments can be produced that do not form part of the lexical segment inventory. A number of properties that can be associated with lexical rules contribute to the coherence and explanatory power of this distinction between lexical and postlexical rules.

Postlexical rules must be distinguished not only from lexical rules but also from phonetic implementation rules. While postlexical rules are phonological rules, and can therefore only manipulate the phonological elements that representations consist of (features, segments, moras, etc.), phonetic implementation rules can only translate the phonological representation into actions of the articulators. The effect of phonetic implementation rules may resemble the effect of phonological rules, and in practice it is not always easy to tell when a rule is phonological.

It might be thought that if lexical rules can be sensitive to morphological information, postlexical rules will be able to refer to syntactic information. There is
some evidence against this assumption. Rather, it would appear that postlexical rules refer to representations that are part of a phonological constituent structure, known as the **prosodic hierarchy**. The theory of Prosodic Phonology, to be discussed in chapter 12, provides an answer to the question of what that constituent structure is and how it might be related to syntactic structure.
9.1 INTRODUCTION

The next three chapters are devoted to phonological elements in the representations of words and sentences that go beyond the strings of consonants and vowels. This chapter introduces the syllable, a constituent that groups segments. Chapter 10 will deal with tones, whether functioning in the representation of words or as part of the intonation, while chapter 11 introduces the foot and word stress.

As you will have come to expect from chapter 3, the allowable segmental compositions of syllables show considerable cross-linguistic similarities, despite their large range of complexity. We will see that the similarity is largely due to the effect of the Sonority Profile, which suggests that the ideal syllable has a burst of acoustic cues (structurally corresponding to the onset) followed by a drawn-out sonorous part (structurally corresponding to the rhyme) (section 9.2). We will next see that between the segments and the syllable an additional level of structure is required. An important role it is meant to fulfil is the representation of segmental duration. There have been two ways in which this intermediate level has been conceptualized. One is the CV-tier, a level of structure between the segments and the syllable specified as C-slots and V-slots, introduced in section 9.3, together with the notions of autosegment and association. The other approach assumes a rhyme constituent which dominates moras, or length units (section 9.4). In either approach, length is represented in terms of the number of units, C-slots and V-slots in the first case and moras in the second. Another important role for the intermediate level of representation is to specify the membership of segments to the subconstituents of the syllable, the onset and rhyme, the latter dividing into peak and coda (earlier introduced in section 3.2). We consider the role of the syllable and syllable membership in phonological generalizations about the distribution and realization of segments (section 9.5), including those that require the existence of ambisyllabic consonants (section 9.6).

9.2 SYLLABIFICATION: THE MAXIMUM ONSET PRINCIPLE

Languages vary greatly in the complexity of their syllables. Some only allow CV, quite a few have (C)V, still others allow a coda, and further complexity is achieved by allowing more than one segment in onset, peak or coda (section 3.2). Independently of this variation, there are striking cross-linguistic similarities. An
intervocalic consonant will occur in the onset of the second syllable, rather than the coda of the first. That is, consonants prefer to form an onset rather than a coda, if they can legitimately do so. The principle responsible for maximizing the onset is (1) (Kahn 1976).

(1) **Maximum Onset Principle (MOP)**
First make the onset as long as it legitimately can be; then form a legitimate coda.

The MOP requires that a string like [tata] should be syllabified [ta.ta], rather than *[tat.a]. And if the language allows [st] onsets, then a string like [asta] will be [a.sta], rather than *[as.ta] or *[ast.a]. Of course, it is essential to keep in mind that, while the MOP is a universal principle, languages differ in the kinds of syllables they allow. For example, Dutch and English allow [st] in the onset, but Spanish does not. Application of (1) to a Spanish string [basta] will therefore yield [bas.ta], not *[ba.sta]. By the same token, German syllabifies extra as [eks.tra], but Dutch as [ek.stra], because Dutch has [st] onsets, but German does not, even though it does have [ft] onsets.

Languages differ in the **syllabification domain.** In West Germanic languages, the syllabification domain tends to be the word, or, rather, the constituent to be introduced in chapter 12 as the phonological word. For instance, the English grey tiles is not homophonous with great isles, the first syllable being [gret] in the first phrase and [greit] in the second, which difference will have an effect on the duration of [ei], in particular. Other languages may have larger domains, like French, which syllabifies across words and where petit ami ‘little friend’ and petit tamis ‘little sieve’ are both [pə.ti.ta.mi], or smaller domains, like Japanese, where certain suffixes will not syllabify with the stem, as shown by /a.ni/ ‘brother’ and /an.i/ ‘easy going’. In the Japanese case, the corresponding phonetic difference is very clear, because the onset nasal is coronal, but the coda nasal more likely to be velar or uvular. That is, /a.ni/ is pronounced [anji], and /an.i/ [anji]. More generally, prefixes in German do not syllabify with the stem, as in Verein [fεrʔain] ‘club’, where Ver- is a prefix.

### 9.2.1 The Sonority Profile

The segmental composition of onsets and codas shows striking similarities across different languages. If you were to bet on the type of consonant that could occupy the second position in a CC onset in some unknown language, your chances of winning would be well served by opting for a glide like [j] or [w]. And if your bet concerned the first consonant, you would be well advised to go for [p], [t] or [k]. Venneman (1972) described these tendencies in terms of a number of syllable ‘laws’. Much of the regularity is captured by the **Sonority Profile**, given in (2).

(2) **Sonority Profile**
The sonority of a syllable increases from the beginning of the syllable onwards, and decreases from the beginning of the peak onwards.
Intuitively, sonority is related to the overall acoustic energy of segments. In (3), the classes of segments that are usually distinguished along this dimension are listed in the order of increasing sonority.

(3) **SONORITY SCALE**

Obstruents – Nasals – Liquids ([l r], etc.) – Glides ([w j], etc.) – Vowels

Thus, the observation is that any onset that reverses the direction of increasing sonority, like [mk-] or [wl-], is less common than one that does not, like [pn-] or [ml-]. Conversely, any rhyme that increases the sonority from left to right, like [-lj], is disfavoured. Violations of the Sonority Profile are indeed rare. Swedish has syllables like [bærj] ‘mountain’, while Dutch has syllables like [vrɪŋ] ‘wring’, where the first consonant is a labiodental approximant. In addition, Clements (1990) observes that in the onset large sonority differences are preferred over small ones, making [pj-] a better onset than [lj-], while in the rhyme small sonority differences are preferred over large ones, making [-j] a better coda than [-t]. Apparently, syllables, like the world in T.S. Eliot’s *The Hollow Men*, prefer to start with a bang and end with a whimper. Because the beginning of the syllable is maximally salient as a result, catching the listener’s attention when it should, the reason for this state of affairs is probably perceptual.

There is an interesting consequence of the observation by Clements (1990) that syllables are marked by a sudden increase in sonority followed by a drawn-out sonorous part. Syllables tend to group in words so that the sonority of the end of one syllable is greater than that of the beginning of the next, favouring a whimper-bang transition over a bang-whimper one. In languages without complex onsets or codas, [al.ka] will therefore be a more likely structure than [ak.la]. This tendency was earlier described as the Syllable Contact Law by Venneman (1972). So also when considering the relation between syllables, the tendency is to maximize the bang at the beginning.

Sonority can be defined in terms of the features introduced in chapter 5 and 6 (Clements 1990). The first four sonority classes of (3) are characterized by the features [±consonantal], [±sonorant], [±approximant]. To distinguish vowels from approximants, syllabicity needs to be called upon, if these are represented in terms of the same features, which in the analysis presented in chapter 6 is the case for [w] and [u], for instance (see also Q145).

**Q80** What do the English words *stigma*, *comrade* and *Daphne* have in common?

**9.3 EXPANDING REPRESENTATIONS: HIERARCHIES AND AUTOSEGMENTS**

In this section, we introduce two ways in which phonological representations go beyond a **linear representation**, i.e. a single string of phonological elements, which was characteristic of earlier theories. There are two ways in which representations are
more elaborate than a single string. One is the hierarchical representation by which segments are dominated by constituents like the syllable. The second is less self-evident. This is the arrangement of different classes of segments as strings that are parallel in time. The next two sections introduce these two enrichments of the representation.

9.3.1 Skeletal slots

There is a lot of evidence that segmental length must be represented separately from the phonological content of segments. A particularly clear example is provided by language games. One of these, Ludikya, reported by Clements (1986) for Luganda, involves a reversal of the order of the syllables in the word. Luganda has both a vowel-length contrast and a contrast between long and short consonants. We indicate long segments by doubling the symbol. The striking thing is that although the syllables move round in this game, the durational structure remains intact, as shown in (4).

(4) Luganda | Ludikya
---|---
mukono | nokomu ‘arm’
mubinikilo | lokinibimu ‘funnel’
baana | naaba ‘children’
ʤuba | bbaju ‘dove’
kiwójolo | löojowwoki ‘butterfly’
kubája | jabadku ‘to work in wood’

In order to express the mutual independence of segmental length and segmental quality, Clements and Keyser (1983) assumed that segments are not immediately associated with their syllables, but are dominated by structural positions, known as (skeletal) slots, which encode segmental length, also known as segmental quantity. Consonants and vowels that are associated with single slots are short, while long vowels and geminate consonants are represented as being doubly linked to two slots. Additionally, they assume that there are two types, a C and a V, where the C represents a syllable margin (onset or coda) and V a syllable peak. That is, CV-slots indicate a segment’s membership of subsyllabic constituents, in addition to providing a representation for segmental quantity.

The CV-slots are dominated by the syllable nodes, each syllable being represented by \( \sigma \). For example, a language with a length contrast for vowels as well as for consonants, like the Dravidian languages Malayalam and Tamil, would have representations like those in (5), which represent the Tamil words [paːtu, paːːtu, paːtu, paːːtu], respectively (Firth 1957; Mohanan 1986: 108). All four words consist of two syllables each and have the same string of segments. However, they differ in their segmental timing structure as expressed by the C-slots and V-slots.

(5) a. CV.CV | b. CV.CV | c. CV.CV | d. CV.CV
---|---|---|---
\( p a \, t u \) | \( p a \, t u \) | \( p a \, t u \) | \( p a \, t u \)
‘to endure’ | ‘to sing’ | ‘enduring’ | ‘a song’
The CV representation allows us to characterize the Luganda word game Ludikya as illustrated in (6). In Ludikya, only part of each syllable is transposed, the string of consonants and vowel segments. The remainder of the representation, the structure above the segments (C-slots and V-slots), is retained. This is shown in (6), where the top line gives the order of the segments in the Luganda word, while the bottom line gives the Ludikya version. Observe how the strings of skeletal slots are identical in the two versions of each word.

\begin{tabular}{ll}
  (6) & b a n a \quad k u b a j a \\
    & \hline \hline \\
    & CVVCV \quad CVVCVCV \\
    & \hline \hline \\
    & n a b a \quad j a b a k u \\
\end{tabular}

In (5) and (6), segments are included in hierarchical structures (section 1.5). A distinction is sometimes made, or implied, between the lines connecting segments and higher constituents and the lines that connect elements of different ranks in the hierarchy, like the slots on the CV-tier and the syllables. Those that connect segments, elements with featural content, to structural constituents are referred to as association lines, the term phonologists use to indicate the structural position of elements with phonological content relative to ‘packaging’ units. Thus, in Luganda, the [j] in ‘to work in wood’ in (6) is associated (also known as ‘linked’) to two C-slots, while in Ludikya it is [k] which has this double association.

### 9.3.2 Autosegments

McCarthy (1985a) used the CV-tier to account for the fact that certain morphemes in Arabic appear to be specified in terms of strings of consonant and vowel positions, referred to as templates. The morphology of the Arabic verb includes a number of derivational morphemes (‘conjugations’), which are not just expressed by particular affixes, but also by different templates. In some cases there is no affix, so that conjugations can differ only in the CV template they have. The verbal root consists only of consonants, usually three. The vowels in such verbal forms represent the third type of morpheme, which corresponds to verbal aspect or voice. In (7), some of these morphemes are listed. A verbal form, then, consists minimally of a verbal root (two or more consonants), a conjugation (a CV template) and a verbal aspect (one or more vowels).

\begin{tabular}{ll}
  (7) & ktb \quad \text{‘write’} \\
    & hq \quad \text{‘be true’} \\
    & CVVCV \quad \text{‘Plain’} \\
    & CVCCVC \quad \text{‘Intensive’} \\
    & CVVCVC \quad \text{‘Influencing’} \\
    & a \quad \text{‘Active Perfective’} \\
    & ui \quad \text{‘Passive Perfective’} \\
\end{tabular}
In (8), the six forms (three conjugations by two aspects) for ‘write’ are given as well as five forms for ‘be true’. The final [a] is 3sg; masc, not shown in (9).\(^1\)

<table>
<thead>
<tr>
<th>Conjugation</th>
<th>Active</th>
<th>Passive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>katab-a</td>
<td>kutib-a</td>
</tr>
<tr>
<td></td>
<td>‘he wrote’</td>
<td>‘it was written’</td>
</tr>
<tr>
<td>Intensive</td>
<td>kattab-a</td>
<td>kuttib-a</td>
</tr>
<tr>
<td></td>
<td>‘he caused</td>
<td>‘it was made</td>
</tr>
<tr>
<td></td>
<td>to write’</td>
<td>to write’</td>
</tr>
<tr>
<td>Influencing</td>
<td>kaatab-a</td>
<td>kuutib-a</td>
</tr>
<tr>
<td></td>
<td>‘he corresponded’</td>
<td>‘he was</td>
</tr>
<tr>
<td></td>
<td></td>
<td>corresponded with’</td>
</tr>
</tbody>
</table>

| Plain         | ħaqq-a          | –                 |
| Intensive     | ħaqqaq-a        | huqqiq-a         |
|               | ‘it is true’    | ‘it has been     |
|               |                | realized’        |
| Influencing   | ħaaqaq-a        | ħuuqiq-a         |
|               | ‘he contested   | ‘his right was   |
|               | sb.’s right     | contested’       |

To derive these forms, we need to assume that the consonants and vowels form separate, parallel segment strings. When there is more than one string of segments in the representation, the segments in each string are known as ‘autonomous’ segments, or *autosegments* for short (Goldsmith 1976). Autosegments were first introduced and motivated for tones, as we will see in the next chapter. Because the strings of consonants and vowels are located on different tiers, consonants can associate with C-slots and vowels to V-slots without getting their association lines crossed or entangled. The C-slots and V-slots themselves form a single string, the CV-tier. In (9), the two segmental tiers are shown above and below the CV-tier. Imagine that between each segment tier and the CV-tier there is a plane across which lie the association lines, as if the CV-tier was the spine of a spiral-bound book with two pages. Association lines between a consonant and non-adjacent C-slots will not be able to interfere with association lines between a vowel and one or more V-slots, because they are drawn on different pages.

<table>
<thead>
<tr>
<th>(9)</th>
<th>k  t  b</th>
<th>k  t  b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C V C V C</td>
<td>C V V C C</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>‘he caused to</td>
<td>‘he was</td>
</tr>
<tr>
<td></td>
<td>write’</td>
<td>corresponded  with’</td>
</tr>
</tbody>
</table>

When a segment associates with two slots, it has *spread* from the first to the second slot if the perception is one whereby you work from left to right through the segmental string and create the associations as you go. For the morpheme ‘be true’ we could either assume an underlying ‘triliteral’ root ([ħqq]) or a ‘biliteral’ one ([ħq]). If we assume that the direction of association is left to right, the form [ħaqqaq-a], for instance, can be accounted for by the spreading of [q] in a biliteral root to all available C-slots, as shown in (10): the left-over C-slots are filled by the spreading of the last consonant. If surface repetitions of the same consonant are produced by the left-to-right spreading of a *rightmost* consonant segment to empty C-slots, as shown in (10), the first two of three surface consonants in verbal and
nominal forms will never be identical. This prediction is by and large correct, a rare exception being [dadan] ‘plaything’ (McCarthy 1985a: 146). Thus, a constraint on the underlying form of Arabic roots is that there should be no adjacent identical consonants. It is akin to the ban on sequences of identical tones, to be discussed in chapter 10, and known as the Obligatory Contour Principle (OCP).

(10) 

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C V C C</td>
<td>C V V C V C</td>
</tr>
<tr>
<td>a</td>
<td>u</td>
</tr>
</tbody>
</table>

‘he realized s.t.’ ‘his right was contested’

Q81 A secret language based on Amharic uses the forms in the second column for the words in the first (McCarthy 1985b). The apostrophe indicates glottalization of the consonant; geminate glottalized consonants are only marked once for glottalization.

<table>
<thead>
<tr>
<th>Amharic</th>
<th>Disguised form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>gainən</td>
</tr>
<tr>
<td>mat’ta</td>
<td>mait’ət’</td>
</tr>
<tr>
<td>kifu</td>
<td>kaifəf</td>
</tr>
<tr>
<td>t’ai’ta</td>
<td>t’ait’ət’</td>
</tr>
<tr>
<td>hed</td>
<td>haidəd</td>
</tr>
<tr>
<td>wəddədə</td>
<td>waidəd</td>
</tr>
<tr>
<td>b</td>
<td>wərk’</td>
</tr>
<tr>
<td>tamara</td>
<td>taim̪rər</td>
</tr>
<tr>
<td>sigara</td>
<td>saigrər</td>
</tr>
<tr>
<td>səkkərə</td>
<td>saikrər</td>
</tr>
<tr>
<td>kəbdəd</td>
<td>kaibdəd</td>
</tr>
<tr>
<td>wifət</td>
<td>waijət</td>
</tr>
</tbody>
</table>

1 What is the difference between the Amharic words and the disguised words?
2 Which segments are preserved in the speech disguised forms, and which are not?
3 What determines the number of (surface) consonants in the speech disguised forms?
4 Give an account of the formation of the speech disguised forms, employing a representation with a skeletal CV-tier. Consider [ai] a single short vowel that associates with a single V-slot.
5 How does your description of this secret language differ from Arabic verbal morphology?
6 Give the representations of the disguised forms for ‘drink’ and ‘gold’.
Scottish Gaelic has a three-way length opposition for vowels, ‘short’, ‘half-long’ and ‘fully long’, as illustrated by [ʃǐːn] ‘we’, [ʃǐːn] ‘venison’, [ʃǐːn] ‘to sing’, or [tʊl] ‘to go’, [uːl] ‘apple’, [suːl] ‘eye’. At most, two different vowel qualities may occur per syllable, as illustrated in [ʃǐːn] ‘John’ and [ʃǐːn] ‘bird’. In addition, diphthongs are either half-long or fully long. Moreover, the distribution of the durations of two intra-syllabic vowel segments is ‘short, short’ (as in [ʃǐːn]) or ‘short, half-long’ (as in [ʃǐːn]), but never ‘half-long, short’ (*[ʃǐːn]) (Ternes 1973: 96). How would you account for these facts?

9.3.3 Unfilled and unassociated slots

The slots in the CV skeleton are structural positions, which in the usual case are dominated by a syllable node and are associated with segments. By assuming that slots don’t have to be associated with any segments and need not be dominated by a syllable, their explanatory role can be extended. Clements and Keyser (1983) call on both these possibilities to account for the alternation between French morpheme-final consonants with Ø, known as liaison, and the phenomenon of h-aspiré.

The French definite article is [la] (masc; sg) or [la] (fem; sg) for the singular, and [le] (pl), as shown in (11a). When the noun begins with a vowel, the vowel of the singular definite article is deleted, while the plural is followed by an apparently inserted [z], as can be seen in (11b). In fact, many words appear to have a consonant that only shows up before a vowel, like [pəti], which has a [t] in [pəti am] ‘little friend’. These potential consonants are called liaison consonants.

(11) Singular  | Plural
--- | ---
\(a\) | \(b\)
la bwa | le bwa  | ‘wood’
la pa | le pa | ‘step’
la karaf | le karaf | ‘carafe’
la nɥi | le nɥi | ‘night’

The loss of the vowel of the singular definite article is effected by elision (12), which deletes the final V-slot of the definite article before V, causing its syllable node to be lost as well.

(12) ELISION

\[
\begin{array}{c}
\text{V} \\
\text{[…]}_{\text{DEF}} \\
\rightarrow \text{Ø} / \_ \_ \text{V}
\end{array}
\]

 Liaison is expressed in (13), which says that an onsetless syllable syllabifies any unsyllabified consonant appearing before the vowel, as shown by the dashed line.

(13) LIAISON

\[
\begin{array}{c}
\sigma \\
C \\
\rightarrow \text{V}
\end{array}
\]

These rules produce the correct results if the singular and plural forms of the article are represented as in (14a,b). The singular will lose its vowel before a V-initial word, while the unsyllabified final consonant in the plural will only be syllabified, and pronounced, before a V-initial word.

\[
\begin{array}{ll}
\text{(14) a. } & \sigma \\
\sigma & C V \\
\sigma & C V C \\
\sigma & \text{a} \\
\sigma & \text{e z}
\end{array}
\]

Certain vowel-initial nouns appear to behave as if they began with a consonant: they take the preconsonantal alternants of the definite article, both in the singular and in the plural, as shown in (15). In order to account for the behaviour of words like ‘hero’ and ‘hatred’, their underlying forms are assumed to begin with an empty C, which is syllabified as an onset C just as any filled C would be. This is shown in (16b), which representation should be compared with (16a).

\[
\begin{array}{ll}
\text{(15) Singular} & \text{Plural} \\
\text{la ero} & \text{le ero} \quad \text{‘hero’} \\
\text{la en} & \text{le en} \quad \text{‘hatred’}
\end{array}
\]

Thus, the deletion of pre-V [a] in (17a), as well as its retention in the pre-C context of (17b), are readily accounted for. Likewise, the difference between the presence of the liaison [z] in ‘the priests’ (18a) and its absence in ‘the heroes’ (18b) is explained by the inability of the unsyllabified [z] to undergo liaison in (18b). This solution accurately expresses the fact that words with h-aspiré behave as if they began with a consonant, even though no surface consonant is observed. Also, it correctly characterizes both the distribution and the identity of the liaison consonant.

\[
\begin{array}{ll}
\text{(17) a. } & \sigma \\
\sigma & C V \\
\sigma & C V C \\
\sigma & \text{i d e} \\
\sigma & \text{e n}
\end{array}
\]

\[
\begin{array}{ll}
\text{(18) a. } & \sigma \\
\sigma & C V \\
\sigma & C V C \\
\sigma & \text{e z} \\
\sigma & \text{a b e z}
\end{array}
\]

9.3.4 Compensatory lengthening

A final argument for the existence of the skeletal tier is the phenomenon of compensatory lengthening. Frequently, the loss of a segment is incomplete in the sense
that the time it took before it was deleted is preserved in a neighbouring segment. The parent language of English and Frisian, which at one time constituted a West Germanic dialect sometimes referred to as Ingwäonic, underwent a process of nasal loss before fricatives within the word. Prior to the loss of the nasal, the vowel before it was short, but a long vowel remains today. In (19), we give reconstructed (hypothetical) forms. The presence of the nasal is attested in Gothic texts, as well as in modern German. (Probably, the nasalization was preserved on the vowel at first, and was lost later. We ignore this in the transcriptions.)

(19) gans ga:s ‘goose’
    fimf fi:f ‘five’
    tanθ ta:θ ‘tooth’
    munθ mu:θ ‘mouth’

The representation with the skeletal tier allows one to express the change as a retiming of the segments, as shown in (20). An additional measure, one which is triggered by the [–cons] segment associated with it, is a change of the vacated C-slot into a V-slot.

(20) Compensatory lengthening can also be recognized in consonants that have come to occupy the syllable peak (syllabic consonants). The second syllable in English bottle [ˈbɒtl] consists purely of the consonant [l]: the release of the plosive [t] takes place through a lowering of the sides of the tongue, the tip retaining its contact (lateral plosion). Assuming that the duration of the syllabic [l] is no different from the duration of [əl], a less common but attested pronunciation, the syllabic consonant could be represented as in (21).

(21) Retiming of segments is a well-known feature of Bantu languages. Class prefixes like [ki] usually only show up in that form before consonant-initial stems. When the stem begins with a vowel, however, the close vowel turns into a glide while the vowel lengthens. Consider the Luganda examples in (22), from Clements (1986).

(22) li+ato lja:to ‘boat’
    mu+iko mwiːko ‘trowel’
    ki+uma kjuːma ‘metal object’
    mi+aka mjaːka ‘year’
If it is assumed that the resulting onset does not consist of two consonants, but of a single, complex consonant which is either labialized (in the case of retimed [u]) or palatalized (in the case of retimed [i]), the process can be described as in (23).

(23) \[ C V + V C V \rightarrow C V + V C V \]

Q83 In Catalan, a word-initial high vowel turns into a glide when preceded by a stressed vowel, as in [kʊnˈtəˈraɾiaˈɾia] `he will tell stories’ and [katəˈlawniʃəɾal] `universal Catalan’ (Mascaró 1970; Hualde 1992). Assuming that the duration of the vowel plus glide is equivalent to that of a short vowel, how could you describe this process in CV phonology?

9.4 MORAS

The skeletal tier fails to account for two phenomena (Hayes 1989a). One is that compensatory lengthening always occurs in the case of segments deleted from the rhyme, never in the case of segments deleted from the onset. This suggests that segments in the rhyme possess something that other segments do not possess. Second, it appears that many languages distinguish syllables on the basis of quantity, a property of syllables which is determined by the number of segments in the rhyme, again to the exclusion of their onset. As for the first objection, consider the fate of the onsets [kn- gn-] in Middle English. When [k g] were lost, and the English syllable no longer admitted onsets that consisted of [−cont] [+nas], there was no compensatory lengthening of any of the other segments in words like knot and gnat, now pronounced [nɔt] and [næt], not *[nɔːt] or *[nɔːt], etc.

Second, the location of word stress frequently appears to be sensitive to the segmental composition of the rhyme, while the number of segments in the onset is typically irrelevant (for some such cases, see Topintzi 2010). In Hawaiian, for instance, stress falls on the last syllable if it contains a long vowel, and on the penultimate syllable if the last is a short vowel, as illustrated by [nəˈna:] `strut’, [ˈnaːna] `for him’ and [loːˈʔihi] `long’. Significantly, a consonant in the rhyme is often counted as if it was a vowel. In Hopi, the stress falls on the first syllable if the rhyme contains a long vowel or a short vowel plus a coda consonant, but on the second syllable if the first syllable contains a short vowel (Jeanne 1982). That is, the consonant in the rhyme can determine the location of the stress in a way that a consonant in the onset cannot.

To capture the difference between segments in the rhyme and segments in the onset, the mora has been proposed as an intermediate level of structure in the rhyme (Hyman 1985; Hayes 1989a). The prosodically ‘active’ status of a segment can then be expressed by giving it moraic status, while ‘inactive’ segments are non-moraic.
A language that has no vowel-length distinction and does not allow a coda will have only monomoraic syllables like (24a). Long vowels look like (24b), while non-moraic and moraic consonants are shown in (24c,d). Geminate consonants must be moraic in order to express their duration. The Italian distinction between [ˈfato] ‘fate’ and [ˈfatːo] ‘fact’ is expressed by representing the singleton consonant as an onset consonant, while the geminate is a consonant that associates with a mora in the first syllable and is a (non-moraic) onset in the second, as shown in (25).

Although it is not uncommon for languages to allow the last syllable of the word to have three moras, languages that more generally allow their rhymes to have three moras are rare. If a language has coda consonants, usually only the first consonant after a short vowel is moraic, while the next consonant or any consonant after a long vowel is non-moraic. Languages that have three degrees of vowel length, like Dinka, and languages that allow a geminate to follow a long vowel have trimoraic syllables. Tamil, as we saw in (5), is an example of a language with geminates after long vowels. The moraic representations of those four structures are given in (26), where (26a) has a monomoraic, (26b,c) a bimoraic and (26d) a trimoraic first syllable.

Trimoraic languages are rare, and it is therefore common for long vowels to be banned before geminates. For instance, Koya words like [aːnd] ‘female power’, [manasuːrku] ‘men’, [meːndulil] ‘back’ are unremarkable in the language, as are short vowels before a geminate, as in [pikːa] ‘cup’. In these words, long vowels appear before a coda consonant, and short vowels before a geminate, but a form like *[puːtːi], in which a long vowel appears before a geminate, is ill-formed. When this configuration arises as a result of morphological concatenation, the vowel is shortened, as shown by /keːːtːoːnda/, which surfaces as [ket.toːnda], where the full stop indicates a syllable boundary. The Koya rhyme thus has maximally two moras, where the second can be either the second half of a long vowel or a consonant;
any additional consonants are non-moraic. Since geminates cannot be non-moraic, they require the vowel before them to be short (Tyler 1969).

**Q84** Give the moraic-cum-syllabic representations of the Koya words [pikːa], [manasu:rku] and [beske] ‘when’.

**Q85** The syllable structure of Tera is CV(X), where X can be V or C. Except at the end of the sentence, word-final [ə] is deleted unless the preceding syllable has a long vowel or when it is immediately preceded by two consonants. Thus, the words in the first column are subject to the rule, but not those in the third (Newman 1970). ([mb] and [nd] are single consonants.)

<table>
<thead>
<tr>
<th>deleted</th>
<th>retained</th>
</tr>
</thead>
<tbody>
<tr>
<td>səɗə</td>
<td>‘snake’</td>
</tr>
<tr>
<td>wurə</td>
<td>‘tomorrow’</td>
</tr>
<tr>
<td>mbukə</td>
<td>‘to throw at’</td>
</tr>
<tr>
<td>pərsə</td>
<td>‘horse’</td>
</tr>
<tr>
<td>me:nə</td>
<td>‘today’</td>
</tr>
<tr>
<td>xəndə</td>
<td>‘nose’</td>
</tr>
</tbody>
</table>

1. What determines if [ə] is deleted?
2. Does Tera have moraic consonants? Motivate your answer.

**Q86** Tagalog has a denominal formation meaning ‘an imitation of’ (Schachter and Otanes 1972). The formation involves the reduplication of the noun and the suffixation of [an], as shown below. Syllable onsets in Tagalog are C or CC, where – at least in the native vocabulary – the C in second position is either [j] or [w]. The syllable may have a coda, which is obligatory in word-final position. In addition to many other consonants, [h] and [ʔ] can occur at the end of the word, but these consonants do not appear word-internally in the coda. Vowel length is in general unpredictable, except in word-final position, where vowels are always long. This can be seen in the forms in both columns. However, vowel length in ‘imitation of’ forms is subject to further restrictions.

<table>
<thead>
<tr>
<th>Base</th>
<th>Derived form</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>ba:ha:j</td>
</tr>
<tr>
<td>b</td>
<td>libru:h</td>
</tr>
<tr>
<td>c</td>
<td>pa:ri:?</td>
</tr>
</tbody>
</table>
Like all phonological constituents, the syllable may have a role in the structural description of phonological generalizations. It is very common for consonants to be pronounced differently depending on their position in the syllable. Thus, as was explained in chapter 7, British English [l] is accompanied by velarization whenever it appears in the coda, as shown in (27). Observe that [lj]- is a legitimate onset, occurring word-initially in a word like lure [ljʊə], for instance.

The need to include syllable structure in English phonological representations was demonstrated by Kahn (1976). He showed that in many cases the context of a rule could only be expressed in segmental terms at the cost of fairly baroque specifications. In the case of 1-VELARIZATION, we would have to state that it applied ‘before all consonants except [j] and at the word end’. Obviously, this approach does not yield a natural class. It would also not be possible to derive the non-velarized allophone from an underlying velarized [h], as in that case we would have to state the context as ‘before all vowels and [j]’. Again, this is not a natural class, since [w] is excluded. Reference to the syllable makes it possible to formulate the context as in (28).

Many other examples can be given of consonants having different allophones depending on whether they appear in the onset or the rhyme. Thus, the Dutch
labial approximant is labiodental [v] in the onset and rounded bilabial [β] in the coda in the varieties spoken in the northern half of the Dutch language area, as shown in (29). Again, neither the generalization ‘in the coda’ nor the generalization ‘in the onset’ can be expressed directly. Since Dutch allows [v:r-] onsets, the onset happens to be equivalent to ‘before a vowel or [r]’, while the coda must be captured by ‘before a consonant other than [r] or a word boundary’. Such structural descriptions clearly look arbitrary and fail to express just what it is that governs the allophony.

(29) In onset In rhyme
[ɛrət] ‘cruel’ [ xe: βt] ‘yawn+3sg’
[ˈkleːʋaŋ] ‘scimitar’ [ˈlyβ̞ tə] ‘lull’

Q87 In Bugotu (Solomon Islands), vowels can be long or short. Any two vowels in either order can form a diphthong, such as [ao], [io], as inferred from the data in Kennedy (2008). It has a reduplicative prefix whose phonological form depends on the base, as illustrated in (1a,b,c,d). The full stop indicates a syllable boundary, and a dash indicates that the unreduplicated form is not used. The language has undominated NOCODA, and a form like *[kal.ka.lu] would therefore be ungrammatical.

(1) a ka.lu – kau.ka.lu ‘to stir up’
ka.ve ‘grandmother’ kae.ka.ve ‘to be old, of woman’
pa.ŋa – paa.pa.ŋa ‘to pant’
lo.po ‘be in a roll’ loo.lo.po ‘to fold’
po.no.ti ‘to close’ poo.po.no.ti ‘to patch’

b koa – koa.koa ‘fornication; to sin’
mee ‘to be foolish’ mee.mee ‘to be importunate’
rei.ŋa ‘to look’ rei.rei.ŋa ‘appearance’

ç a.θo ‘cord, rope’ a.θoa.θo ‘to signal’
i.li ‘to totter’ i.lii.li ‘to be drunk’
u.do.lu ‘to be gathered’ u.dou.do.lu (no gloss available)

d ou – ou.ou ‘to cough’
ao – ao.ao ‘crow, bird’
iu ‘dog’ (mane) iu-iu ‘loafer’

The prefix has a segmentless structure of two moras which is filled with a copy of the segments from the base, left to right. Unparsed segments are deleted. Observe that in this problem a morphological operation involves the addition of a specific phonological structure which packages segments. This frequent phenomenon is known as prosodic morphology (McCarthy 1998).
The underlying form of [kau.ka.lu] is thus as in (2a) and that of [poo.po.no.ti] as in (2b).

(2) a  
\[ \begin{array}{c}
\text{kalu + k} \\
\sigma \\
\mu \\
\end{array} \]

b  
\[ \begin{array}{c}
\text{ponoti + p} \\
\sigma \\
\mu \\
\end{array} \]

1. Show the autosegmental surface representation of [kau.ka.lu].
2. \(\text{Dep}(\sigma)\) bans the addition of syllable nodes, \(\text{Max}(\mu)\) bans the deletion of moras, and \(\text{Onset}\) requires that syllables have onsets. Draw a tableau with these three constraints and (2a) as input. Assume moraic-cum-syllabic representations of [kau.ka.lu], *[au.kal.u], *[a.lu.ka.lu], *[ka.kal.u] and *[ka.lu.ka.lu] as candidate outputs. Do these constraints have to be ranked to make [kau.ka.lu] the winner?
3. The tableau you drew up for question 2 explains why [a.\(\theta\)a.\(\theta\)a] is better than *[a.\(\theta\)a.\(\theta\)a] and *[a.\(\theta\)a.\(\theta\)a.\(\theta\)a], but not why it is better than *[\(\theta\)o.a.\(\theta\)o]. Show these four candidates in a tableau and propose a constraint which is violated by *[\(\theta\)oo.a.\(\theta\)o] but not by [a.\(\theta\)oa.\(\theta\)a]. Does your constraint have to be ranked with the other three?
4. What would be the reduplicated form of a hypothetical loanword [ko.kao]?

9.6 POST-MOP SYLLABIFICATION RULES

The MOP (section 9.2), together with the specification of a syllabification domain within which it is applicable, suffices to explain the syllable structure of many languages. Other languages have additional syllabification rules. The motivation for these additional rules is that consonants may behave as if they are in the coda according to one generalization, while at the same time acting as onset consonants of the following syllable according to another generalization. There have been two syllable-based approaches to this situation. One is to redo the initial MOP-driven syllabification after a derivation or cliticization, so that what appeared as a coda consonant at one level of representation appears as an onset consonant at another, a description known as \textbf{resyllabification}. If the first syllabification was V.CCV, as shown in (30b), the second would be (30c). Other descriptions will go on the assumption that the MOP is a persistent principle and that additional syllabification rules must respect existing syllabification. The new syllabification is ‘added’ to the old, and as a result a consonant may simultaneously belong to the coda of one syllable and to the onset of the next: such consonants are said to be \textbf{ambisyllabic}. As stressed by Kahn, the latter position is more restrictive, in that it limits the number
of possible syllabifications. In the hypothetical situation (30), the syllabification required by the MOP would be V.CCV, and only the first C can be ambisyllabic (30a), since the No Crossing Constraint rules out the ambisyllabic of a second C. By contrast, in a resyllabification analysis, there is no principled limit to the number of consonants that can be resyllabified. Starting out with (30b), we could change that into (30c), but there is no reason why a further resyllabification could not change that to VCC.V (not shown). In this section, we will argue that ambisyllabic can be motivated on the basis of Kahn’s description of (American) English flapping.

9.6.1 Ambisyllabic in English

English has two post-MOP syllabification rules which create ambisyllabic consonants. Liaison applies across word boundaries so as to cause a word-final consonant to be in the onset of a following vowel-initial word. Since it doesn’t thereby lose its association as a coda consonant, an ambisyllabic consonants is created. To return to the example in section 9.2, /t/ in great isles is both coda in great and onset in isles, while /t/ in grey tiles is exclusively an onset consonant. Rule (31) only applies across word boundaries, because word-internally onsets will have been created by the MOP. Some more examples are given in (32). After the application of Liaison, the word-initial vowel can no longer be preceded by a glottal stop.

The other rule creating ambisyllabic consonants is Right Capture, which causes the first consonant of an unstressed syllable to serve as a (final) coda consonant of a preceding syllable, as formulated in Gussenhoven (1986). That is, it is a syllabification rule which applies within the domain of the foot. The English foot is left-dominant, in which the weak syllable(s) typically have [ə i o] or, especially word-finally, [i oʊ]. Note that other vowels are strong and, therefore, define a foot (see chapter 11). Foot-based right capture, given in (33), spreads an onset consonant to a preceding syllable, where it forms part of its coda.
As a result of (33), the [p] in a word like happy will be the final consonant of the first syllable and the first consonant of the second syllable. As in the case of MOP, we must make the assumption that the resultant coda is well-formed in English. That is, right capture cannot create a syllable *[bedf]* in *[ˈbedfɔːrd]* Bedford, even though this word is a single foot. Examples of words in which consonants become ambi-syllabic by (33) are given in (34). In (34a), the rule applies after an open syllable, in (34b,c) it applies after a closed syllable, while (34d) shows that the left-hand syllable may be unstressed.

(34)  

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>ˈsɪti</td>
<td>[t]</td>
<td>city</td>
</tr>
<tr>
<td>b</td>
<td>ˈkænsərt</td>
<td>[s]</td>
<td>concert</td>
</tr>
<tr>
<td>c</td>
<td>ˈkantri</td>
<td>[t]</td>
<td>country</td>
</tr>
<tr>
<td>d</td>
<td>əˈspərɑːɡəs</td>
<td>[r]. [ɡ]</td>
<td>asparagus</td>
</tr>
</tbody>
</table>

As pointed out by Kahn (1976), American English flapping, the rule that causes certain instances of intervocalic [t d] to be pronounced as an alveolar flap, provides evidence that ambi-syllabic consonants are a natural class. Flapping applies in what are at first sight two entirely unrelated contexts. In (35), the contexts in which the rule applies are compared with a context in which it does not. The curious generalization that emerges is that flapping applies when the right-hand vowel is in the next word, or is in a weak syllable in the same word.

(35)  

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>flapping applies</td>
<td>No flapping</td>
<td>flapping applies</td>
<td></td>
</tr>
<tr>
<td>[ˈlɛtər] later</td>
<td>[ˈlɛtɛks] latex</td>
<td>[ˈlɛt əks] late ex</td>
<td></td>
</tr>
<tr>
<td>[ˈsɪti] city</td>
<td>[ˈsætərɪ] satire</td>
<td>[ˈnæt oʊl] night owl</td>
<td></td>
</tr>
</tbody>
</table>

The unifying element of the contexts in which flapping applies is ambi-syllabic-ity. Liaison and right capture, which can both be motivated on independent grounds (cf. also Rubach 1996), provide the appropriate representations for rule (36), which simply says 'Flap [t d] when they are ambi-syllabic'.

(36)  

\[
\text{FLAPPING } [t, d] \rightarrow [r]/[\text{cons}]
\]

It has been claimed that languages do not contrast ambi-syllabic and geminate consonants, which suggests that they be represented identically within words (van der Hulst 1985). Accordingly, an ambi-syllabic consonant has an association with the last mora of the syllable it closes and with the onset of the next, as shown in (37a) for Old English and (37b) for city.

(37)  

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>σ</td>
<td>σ</td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>μ</td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>μ</td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td>o u l d</td>
<td>g l</td>
<td>t</td>
</tr>
<tr>
<td>b</td>
<td>σ</td>
<td>σ</td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>μ</td>
<td>μ</td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td>σ</td>
<td>σ</td>
<td>μ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>s</td>
<td>i</td>
<td>t</td>
</tr>
</tbody>
</table>
There are three alternative solutions that make no use of amabisyllabicity. First, in a description without syllabic structure, we would arrive at (38), after Kahn (1976), where the angled brackets indicate mutually dependent terms.

\[
\text{(38) FLAPPING } [t\ d] \rightarrow [r] / [-\text{cons}] \quad \langle\#\rangle_a \quad +\text{syl} \quad \langle \text{−stress} \rangle_b
\]

Condition: if not \(a\), then \(b\)

In addition to the awkward condition ‘within the word, the following vowel must be unstressed’, we also need to stipulate that in slow, deliberate speech styles, in which rules like liaison do not apply, word-final \([t\ d]\) will not be flapped. This will make it necessary to add a second condition: ‘In unconnected speech, not \(a\).’ The second alternative is a resyllabification account. This theory will have to assume that flapping applies to coda consonants (Selkirk 1982). In order for this account to go through, there must be no liaison (in order to keep final \([t\ d]\) in coda position), while right capture must be replaced with a rule that delinks the initial onset consonant of a weak syllable and moves it into the coda of a preceding syllable, so that get \$ Anne, get \$ it and get\$ing are syllabified as indicated by \$. The two-stage syllabification will have to face a number of surface facts. One is that real coda consonants, like \([d]\) in Too bad! or \([t]\) in The lot, are not flapped, but are in fact frequently unreleased. Second, \([t\ d]\) in target flapping contexts are neutralized (cf. bidder – bitter), but contrast in the coda (bid – bit). Third, the claim that weak syllables in words like city or sadder are onsetless is difficult to square with the general finding that real onsetless syllables may be pronounced with a preceding \([\?]\). Moreover, as was observed above, there is the general drawback that the resyllabification option will have to be constrained by some independent principle, since it would otherwise freely allow changes of onsets into codas and vice versa. The third solution to the definition of the context of flapping is one which ties the data to the domain of the foot. This is a possible account, but with a twist. If flapping is the default implementation of \([t\ d]\) rather than the special case, the complement of Kahn’s generalization can be given. At the beginning of the word and the foot, \(/t/\) is realized as \([t^h]\), as in tomorrow, retain, and in the syllable coda as an unreleased \([t^\prime]\), as right, atlas. Elsewhere, it is flapped (Harris 2004).

Q88 In some dialects of Spanish (Harris 1983), \([r]\) and \([l]\) are realized as \([j]\) in certain positions, as illustrated in (1).

<table>
<thead>
<tr>
<th>Word</th>
<th>Realization</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>revolver</td>
<td>rev(\text{j})vej</td>
<td>‘revolver’</td>
</tr>
<tr>
<td>carta</td>
<td>kaj(\text{ta})</td>
<td>‘card’</td>
</tr>
<tr>
<td>papel</td>
<td>pape(\text{j})</td>
<td>‘paper’</td>
</tr>
<tr>
<td>calor</td>
<td>kal(\text{o})j</td>
<td>‘heat’</td>
</tr>
</tbody>
</table>

Give two rule-based accounts, one with and one without reference to the syllable, and argue that the first is superior.
Q89  British English (RP) /r/ can be realized either as a flap [ɾ] or an approximant [ɹ] in the expressions in columns I and III, but in those in column II only an approximant realization is available (Rubach 1996).

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>courage</td>
<td>courageous</td>
<td>for example</td>
</tr>
<tr>
<td>very</td>
<td>reduce</td>
<td>for instance</td>
</tr>
<tr>
<td>baron</td>
<td>red</td>
<td>the other end</td>
</tr>
<tr>
<td>laurel</td>
<td>bright</td>
<td></td>
</tr>
</tbody>
</table>

1. Can these data be described on an assumption of ambisyllabicity?
2. Can these data be described in terms of foot structure?

Q90  The structure of the syllable rhyme of Brazilian Portuguese (BP) is V(X), where X can be C or V. BP has the nasals [m n ɲ]. Prepalatal [ɲ] distinguishes itself from the other nasals in a number of respects. First, while [m n] can appear word-initially, as shown in (1a,b), [ɲ] cannot, as shown in (1c), where V stands for any vowel (Wetzels 1997b).

(1)  
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>‘mɔvel</td>
</tr>
<tr>
<td>b</td>
<td>‘nɔrti</td>
</tr>
<tr>
<td>c</td>
<td>*ɲV. .</td>
</tr>
</tbody>
</table>

Second, although diphthongs are freely formed out of consecutive vowels before other nasals (2a,b), they do not occur before [ɲ]. Instead, the two vowels are divided over two syllables (2c,d).

(2)  
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>‘rei.nu</td>
</tr>
<tr>
<td>b</td>
<td>an.’dai.mu</td>
</tr>
<tr>
<td>c</td>
<td>fu.’iɲa</td>
</tr>
<tr>
<td>d</td>
<td>ra’hja</td>
</tr>
</tbody>
</table>

Third, while [m n] tolerate a consonant immediately before them (3a,b,c,d), [ɲ] does not (3e).

(3)  
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>a’dornu</td>
</tr>
<tr>
<td>b</td>
<td>‘arma</td>
</tr>
<tr>
<td>c</td>
<td>‘alma</td>
</tr>
<tr>
<td>d</td>
<td>vulne’ravel</td>
</tr>
<tr>
<td>e</td>
<td>*VŋɲV, *VlɲV</td>
</tr>
</tbody>
</table>
The three differences between \([n]\) on the one hand and \([m\ n]\) on the other can be explained if \([n]\) is incorporated differently in syllable structure from \([m\ n]\).

1. What is this structural difference?

BP has a rule of vowel nasalization which causes a stressed vowel to be nasalized before a nasal, as shown in (4a,b). Unstressed vowels remain oral before nasal consonants, as shown in (4c,d). However, before \([n]\), vowels are always nasalized, as shown in (4e,f).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>‘sinu’</td>
<td>‘bell’</td>
</tr>
<tr>
<td>b</td>
<td>‘kâma’</td>
<td>‘bed’</td>
</tr>
<tr>
<td>c</td>
<td>bo ‘neka’</td>
<td>‘doll’</td>
</tr>
<tr>
<td>d</td>
<td>kuma ‘ri’</td>
<td>‘chilli’</td>
</tr>
<tr>
<td>e</td>
<td>a ‘râña’</td>
<td>‘spider’</td>
</tr>
<tr>
<td>f</td>
<td>di’ñeiru</td>
<td>‘money’</td>
</tr>
</tbody>
</table>

It is possible to account for the nasalization of stressed vowels before \([m\ n]\) and the lack of nasalization of unstressed vowels before \([m\ n]\) by assuming that in stressed syllables consonants like \([m\ n]\) are incorporated differently in syllable structure than in unstressed syllables.

2. What is that assumption?

3. What is under that assumption the correct generalization for vowel nasalization?

4. How does your solution account for the fact that before \([n]\) both stressed and unstressed vowels are nasalized?

### 9.7 Conclusion

In this chapter we have seen that segmental duration can be accounted for by assuming that there is an independent tier of skeletal slots which constitutes an intermediate level of structure between the segments and the syllable. In addition, the elements on this tier serve to define the morphemes for verb aspect in Arabic. By allowing C-slots to be unsyllabified, it also appeared possible to account for French liaison, while an empty C-slot could be used to account for h-aspiré in French. However, a distinction must be made between segments that determine the status of the syllable with respect to foot assignment and cause compensatory durational effects, on the one hand, and those that do not, on the other. The mora serves to represent the special prosodic status of these segments. The moraic theory predicts that compensatory effects and sensitivity to stress rules go hand in hand in any one language. It is usually assumed that the mora tier should replace, rather than
supplement, the CV-tier. We have also seen that the syllable plays an active role in the SDs of phonological rules, which often appear sensitive to whether a consonant is in the coda or in the onset. Additionally, it has been shown that ambisyllability is a form of representation required for the description of a number of processes.

Of course, the advantages of allowing enriched representations like ambisyllability need to be weighed up against a sparser inventory of representations, which will place more of the descriptive burden on the availability of different levels of representation, as in Selkirk’s (1982) account of American English flapping. However, in the next chapter, we will see how multiple linking for a third kind of segments besides vowels and consonants, tones, has brought many benefits in its wake. Many phenomena discussed there in fact lay at the origin of the idea of autosegmental phonology and multiple linking.

NOTE

1 The expected form *[ḥaq-a] ‘it is true’ is ruled out because of a general process deleting vowels in this context.


## 10.1 INTRODUCTION

In addition to consonants and vowels, phonological representations contain tones (section 2.2.5). In more than half of the languages in the world, tones are used along with vowels and consonants to keep morphemes apart, whether major-class words or function words, in the same way that languages use vowels and consonants for this purpose. Standard Mandarin, for instance, has four lexically contrastive tone patterns, as shown by the minimal quadruplet in (1a,b,c,d). In (1e), a fifth word is given, which has the same tone pattern as (1d) but has a different consonant, minimally contrasting with (1d).

\[
\text{(1a)} \quad \text{baa} \quad \text{eight'} \\
\text{(1b)} \quad \text{baa} \quad \text{pull up (grass')} \\
\text{(1c)} \quad \text{baa} \quad \text{grasp'} \\
\text{(1d)} \quad \text{baa} \quad \text{father'} \\
\text{(1e)} \quad \text{maa} \quad \text{scold'}
\]

Equally, tones may have morphological roles, as when they express definiteness, number or gender on nouns; person, aspect or tense on verbs; or morphological class. Whereas Standard Mandarin only has lexical tones, many other languages additionally, or in a case like Somali even exclusively, have such grammatical tones (Hyman 2016). By the side of lexical tones, many Otomanguean languages have more or less regular correspondences between tones on the verb and their morphological class. For instance, Zenzontepec Chatino has mid tone on the potential and progressive forms of /-sôʔ/ ‘pick’, a bound morpheme representative of verb class A2, but high tone in the completive, /-sóʔ/. And when used in verbs, the eight tones of Iau indicate aspectual meanings, like ‘totality of action-durative’, as in [taiʰ] ‘be pulling off’ and ‘resultative-durative’, as in [tai] ‘have been pulled off’, differently from their lexical functions in nouns, like [teʰ] ‘flooring’ and [te³] ‘mosquito’ (Bateman 1990). Languages that use at least some tones either lexically or grammatically are known as tone languages. They are mainly found in Asia, particularly in South East Asia, as well as in Africa and the Americas. A third function of tones is to add a discourse meaning to the expression or to indicate prosodic phrasing. Most Austronesian and Indo-European languages only have tones with these functions. In English, for instance, the expression Jonathan may sound like an answer to a question in (2a), like a question in (2b), and like either a reminder or a question in (2c).

\[
\text{(2a)} \quad \text{Jonathan} \\
\text{(2b)} \quad \text{Jonathan} \\
\text{(2c)} \quad \text{Jonathan}
\]
A phrasing function of tone occurs in an English sentence like *Once we’re in China, we can practise our Chinese*, where the last syllable of *China* is likely to have a high tone indicating the boundary. When tones function to signal discoursal meanings or phrasing, they are **intonational tones**. The term ‘intonation language’ is not common, because it would incorrectly imply that tone languages have no intonational tones. Many tone languages have intonational boundary tones, for instance.

While there have been a number of proposals for the featural analysis of tones, like \([\pm \text{highpitch}]\) to characterize high and low tones (H vs L) (see Bao 1999; Hyman 2010), in this chapter we refer to tones as integral segments (H M L), where M is mid tone. Importantly, tones are autosegments. They are arranged on a separate tier from the segmental tier (or tiers, if vowels and consonants appear on separate tiers, as in Arabic verbal morphology, section 9.3). Goldsmith (1976) showed that the tonal patterns of the words of many African languages can be insightfully described autosegmentally. We repeat his point that a linear representation is inadequate for explaining those data (section 10.2) and reproduce some of his arguments in sections 10.3 and 10.4. First, when languages have a small set of lexical tone configurations, or **word melodies**, distributed over words with different numbers of syllables, it makes sense to keep the tones and the segments as separate strings in the lexicon and associate the tones with structural positions according to a general algorithm. A second argument for having a separate tone tier is based on the fact that vowels may be deleted but their tone preserved (**tone stability**), while a third argument is based on the existence of morphemes that consist only of tone (**tonal morphemes**). Section 10.5 presents the **Obligatory Contour Principle** (OCP), while section 10.6 introduces the notion of **accent** as a placeholder for tonal configurations. Section 10.7 deals with the phonetic implementation of tones, distinguishing their **scaling** from their **alignment**. The **privative representation** of tone is briefly discussed in section 10.8, and section 10.9 discusses consonantal **f0 perturbations** and some of their effects on tone distributions.

### Q91 What is wrong with the following statements?

1. In a tone language, a change in pitch can change the meaning.
2. In a tone language, a change in pitch can change the meaning of a word.
3. German is a front rounded vowel language.

### 10.2 THE INADEQUACY OF A LINEAR MODEL

Until the 1970s, it was assumed that segments form a single layer of structure, a single string of self-contained feature matrices, whereby tones were incorporated as pitch features in vowel matrices. This linear conception of segmental structure makes it impossible to represent aspects of pronunciation that characterize more
than one segment as a single feature. Languages in fact frequently treat particular aspects of pronunciation, notably tone, nasality and tongue body features, as if they belonged to longer sequences of segments. For instance, if a language were to have words whose syllables were either all high-toned or all low-toned, then it would make sense to say that each word of the language either had the feature [+hightone] (’H’) or [–hightone] (’L’). In the SPE representation of feature matrices, we would not be able to express this generalization: every vowel would have to be specified for [±hightone]. If we did that, the question arises why the words of this language have either consistent strings of [+hightone] specifications or consistent strings of [–hightone] specifications, without ever mixing these in the same word. We could of course construct a grammar that demands this consistency, through the introduction of some constraint that requires that all the segments in a word have the same specification for feature X. However, the same effect would be obtained if feature X was part of the underlying representation of the word and some algorithm distributed it over all the relevant segments. This would allow us to express directly that words are either high-toned or low-toned. That is, we could represent the tonal features, making up the tones, and the other features, making up the other segments, as two separate strings and then allow different numbers of segments in the two strings, such that a word might consist of six vowels, say, but of only one tone. We show the difference for a hypothetical high-toned word [ta ta] in (3): (3a) gives the linear situation adopted in SPE (where [a] stands for the rest of the distinctive features in that matrix), and (3b), or equivalently (3c), gives the representation with two parallel strings. Equally, the tonal string might have more tones than the segmental string has vowels, opening up the possibility of more than one tone characterizing one vowel.

(3)
\[
\text{a} \quad \text{t} \quad \begin{cases} a \\
+\text{hightone} \end{cases} \quad \text{t} \quad \begin{cases} a \\
+\text{hightone} \end{cases} \quad \text{b} \quad \text{ta ta} \quad \begin{cases} c \\
+\text{hightone} \end{cases} \quad \text{H}
\]

Goldsmith (1976) characterized the SPE position as a theory that adheres to the Absolute Slicing Hypothesis: the phonological representation of a word is given by making a number of clean cuts along its time axis, each slice so produced being a segment. The representation which he argued for instead is one in which the cuts can be made for subsets of features, such that slices made for the features specifying vowels and consonants leave the features of tones unaffected, and vice versa.

10.3 WORD MELODIES

If, regardless of the number of syllables, the words of a language were either entirely high-toned or entirely low-toned, a situation briefly considered in the previous section, it would obviously be economical to specify each word once either for low tone or for high tone. Put differently, we would wish to say that the language had two word melodies, H and L, and that each word had either the one or the other.
Etung presents a somewhat more complex case (Edmondson & Bendor-Samuel 1966). A syllable in an Etung word has high, low, falling or rising tone. Importantly, however, there are certain constraints on the distribution of the tone types: not all of them can appear on all syllables. The IPA accent symbols used in the transcription of tone are given in the first column of (4), and the tonal representations in the third.

(4) á  high tone        H
    à  low tone         L
    â  falling tone     HL
    â  rising tone      LH
    â  falling-rising    HLH
    â  rising-falling    LHL

In (5a,b,c), we list actual words of one syllable, two syllables and three syllables, respectively. Notice that level tones (i.e. a high tone or a low tone) may occur on all syllables, but that contour tones (falling tone or rising tone) only appear on monosyllabic words and on the final syllable of disyllabic words. Moreover, if a syllable has the same tone as the preceding syllable, all following syllables have that tone. Thus, the words in (5a,b,c) are all fine, while those in (5d,e,f) are all bad. In (5d), we see that contour tones must be final, in (5e) we see that contour tones cannot appear in words of more than two syllables, while in (5f) we see that in words of three syllables, the first two cannot have the same tone if the third has a different tone.

(5) a  monosyllabic words  kpá, kpè, ná, nô
    b  disyllabic words    óbá, èkát, óbô, ódá, òbô
    c  trisyllabic words   ékúé, éjúři, ók胼à, édimbá
    d  disyllabic          *âbó, *âbó
    e  trisyllabic         *âbòmbá, *âbòmbá, *âbòmbá
    f  trisyllabic         *âbòmbá, *âbòmbá

Leben (1973) argued that such distributional facts can be captured by assuming that tones are not chosen per syllable, but per word, and that, moreover, these tone patterns are independent of the vowels and consonants. These Etung tone patterns are given in (6). Importantly, the tonal tier and the segmental tier are unassociated in the lexicon: the point is that all and only the well-formed lexical patterns of (5) can be derived by an association algorithm, so that we account for the ungrammaticality of the forms in (5d,e,f). To achieve this, Goldsmith (1976) provided the Association Convention (7). The phonological constituent with which a tone associates, the tone-bearing unit (TBU), is either the mora or the syllable. In Etung, the TBU is the syllable. For typographical convenience, a tone's association line is with the vowel of the syllable concerned.

(6) A word has one of the six patterns L, H, LH, HL, LHL, HLH.

(7) Association Convention
    a  Associate tones and TBUs left to right, one to one.
    b  Associate left-over TBUs with the last tone.
    c  Associate left-over tones with the last TBU.
In (8), a number of examples have been worked out. In (8a), the number of tones and the number of TBUs is equal, and one-to-one association is all we need to do (clause (7a)). In (8b,c), there are fewer tones than there are syllables. According to clause (7b), we must now associate the last tone with the left-over syllable(s), an operation known as spreading (see section 9.3.2). Notice that this causes the same tone to be associated with more than one TBU. In (8d), there are more tones than there are TBUs. The last TBU receives the left-over tone (clause (7c)), causing this TBU to be associated with a contour tone.

(8) Lexical representations

```
(7a)
\begin{tabular}{l|l|l|l|l}
   & a & b & c & d \\
\hline
edimba & HLH & LH & H & LHL \\
\end{tabular}

(7b)
\begin{tabular}{l|l|l|l|l}
   & a & b & c & d \\
\hline
edimba & bisone & ekue & obo \\
\hline
H & L & H & L & H \\
\end{tabular}
```

(7c)

```
\begin{tabular}{l|l|l|l|l}
   & a & b & c & d \\
\hline
edimba & bisone & ekue & obo \\
\hline
\hline
L & H & H \\
\end{tabular}
```

Etung has no falling-rising or rising-falling words of the type *ná*. Apparently, there is a limit to the number of tones that can associate with a TBU in this language. This No Crowding Constraint, which appears to be quite common in languages generally, is given in (9).

(9) No Crowding Constraint: 

```
\begin{tabular}{l}
\hline
\hline
\end{tabular}
```

The above description accounts elegantly for the distribution of the four surface tone types (high, low, falling, rising) of Etung. If we were to elevate those surface tone types to features inside a segmental feature matrix, a description would result that would consist of a series of arbitrary facts. In (10), this alternative is made explicit. This cumbersome description should be compared to the simple autosegmental alternative consisting of (6), (7) and (9).

(10) a A syllable has one of the four tones High, Low, Fall, Rise.

b Fall and Rise never occur on a non-final syllable.

c Words of more than two syllables never have a Fall or a Rise.

d Words of more than two syllables never have the same tone on the first n syllables if the nth + 1 syllable has a different tone, where n > 1.

The TBU may be defined differently for different languages. In some languages the syllable is the TBU, so that long vowels count as one TBU. As we will see in some of
the questions, other languages have the sonorant mora as the TBU, in which case long vowels are two TBUs, while sonorant consonants in the syllable rhyme are also TBUs.

**Q92 The majority of Mende words have the tone patterns illustrated below:**

- **one syllable:** kó ‘war’, kpà ‘debt’, mbù ‘owl’, mbà ‘rice’, mbà ‘companion’
- **two syllables:** pélé ‘house’, njáhá ‘woman’, ñgilà ‘dog’, fändé ‘cotton’, bèlè ‘trousers’
- **three syllables:** kpàkàlì ‘tripod chair’, félà mà ‘junction’, ndàvú lá ‘sling’, nikíli ‘groundnut’, háwáná ‘waistline’

1. What are the underlying tone patterns of Mende?
2. Does Mende, like Etung, obey the No Crowding Constraint?
3. Give the tonal representations of the words for ‘woman’ and ‘tripod chair’.

### 10.3.1 Language-specific associations

The Association Convention (7) is not an obligatory procedure in tone languages. For one thing, there may be language-particular rules which override it. In Kikuyu, the association proceeds from the second TBU in the word, rather than the first. The word in (12), which means ‘way of releasing oneself quickly’, is made up of the morphemes in (11). The verb root extension [aŋ] ‘quickly’ is toneless. The association of the tones contributed by the various morphemes proceeds as predicted by (7), with the last H spreading to the left-over TBUs, except that instead of associating with the first TBU, the first tone associates with the second TBU. Subsequently, the first empty TBU is associated with the first tone, as shown in (12) (Clements and Ford 1979).

(11) mo érek+aŋ erie
L H L H
CLASS PREFIX REFLEXIVE allow+quickly NOUN SUFFIX

(12) mo e̱nx ka nge ri e
L H L H

**Q93 The Kikuyu words for ‘firewood’ and ‘tree’ are [ròkò] and [mòtò].**

1. What are the underlying tone patterns of these words?
2. In Tharaka, a related Bantu language spoken in Kenya, these words have the same underlying tone patterns but are realized as [ròkò] and [mòtò]. How would you describe this difference between Kikuyu and Tharaka?
Thus, languages may have specific constraints on associations of particular tones. For instance, Hyman and Ngunga (1994) show that the TBU with which a tone associates in Ciyaq verbal stems may depend on properties like ‘positive’ (H on the last mora), ‘negative’ (H on the second mora) or ‘infinitive’ (H on the first mora, with spreading). Even in languages like Etung, whose word tonal patterns are apparently governed by the Association Convention (7), there may be exceptional patterns that cannot be so accounted for. Likewise, there is no universal requirement that tones should always spread to empty TBUs. Pulleyblank (1986) shows that in Tiv, non-initial TBUs may remain unspecified for tone until a point in the derivation when they receive a default L. Neither need it be the case that left-over tones always form contours on the last TBU. Languages may not tolerate contouring at all, obeying an even more restrictive constraint than (9).

Q94 Zulu (Laughren 1984) obeys a more restrictive constraint than the No Crowding Constraint (9), since no contour tones may appear on short vowels at all. Zulu words behave either like Kikuyu words or like Etung words. That is, every word is lexically specified as to whether association starts with the first or the second TBU. In Natal Coast Zulu, the word for ‘chiefs’ consists of two morphemes, /ama/ and /kosi/, both of which have underlying HL. The first morpheme is specified for association to start with the first TBU, the second for association to start with the second TBU. After this initial association, the rest of the derivation is taken care of by the Association Convention in (7).

1 How will the word for ‘chiefs’ be realized?
2 Why is the realization not [ámà kòsi]?

While the Association Convention in (7) may well represent an unmarked procedure, evidently languages introduce language-particular instructions or have representations that deviate from its predictions. The only uncontroversially universal aspect of autosegmental representations is the No Crossing Constraint, given in (13). It is an inalienable feature of the autosegmental representation: if it did not hold, the order of the autosegments on different tiers could change relative to the orders on other tiers. The constraint in (13) is there to prevent such order changes.

(13) No Crossing Constraint: Association lines do not cross.
In this section, the autosegmental representation of tones was argued for on the basis of distributional restrictions that they are subject to in various languages.

Q95  West Greenlandic, a language that only has intonational boundary tones, has a phrase-final melody HL and an utterance-final H, which come together as HLH at the utterance end. The timing of these boundary tones is illustrated in the following data (Rischel 1974; Nagano-Madsen 1993; preceding unmarked syllables have default tone, realized as middish pitch).

a  akìvârá       ‘I answered him’
b  atâːsíq        ‘one’
c  akìvâː       ‘he answered him’
d  úvânjâ        ‘I’
e  uvânjâlú      ‘and I’
f  uvaŋâtːâː  ‘I, too’

1 What is the TBU in West Greenlandic?
2 Show the structure of the expressions in a, b, c. Show long vowels with double symbols (i.e. [aː] as [aa]).
3 What is the direction of association of the tones?

Yes-no questions end in the same melody but are distinguished from declaratives by the addition of a mora to the final syllable.

<table>
<thead>
<tr>
<th>Declarative</th>
<th>Interrogative</th>
</tr>
</thead>
<tbody>
<tr>
<td>g takusâːná</td>
<td>takusâːnáː     ‘did he see me?’</td>
</tr>
<tr>
<td>h takusâːŋuk</td>
<td>takusâːŋuk      ‘did you see him?’</td>
</tr>
<tr>
<td>i tsiġuusâː</td>
<td>tsiġuusâːː      ‘did he take that?’</td>
</tr>
<tr>
<td>j a pièrâi</td>
<td>a pièrâːi       ‘did he ask them?’</td>
</tr>
</tbody>
</table>

4 Give the tonal representations of the declarative and interrogative forms of ‘he asked them’.
5 The penultimate vowel of the interrogative form in (g) is fully high-toned. What extra assumption must be made to account for this fact?
6 What is the direction of association of the elements in a diphthong?

10.4  TONAL STABILITY AND TONAL MORPHEMES

Sometimes, a vowel is deleted, but the tone it had shows up on an adjacent vowel. In (14), some examples are given from Etsako, in which a nonhigh vowel is deleted before another vowel (Elimelech 1976). When the final vowel of the stem is deleted,
its tone shows up on the next vowel, where it creates a contour. If the tone of the deleted vowel is the same as the tone on the following vowel, however, nothing happens. In Etsako, reduplication of the noun has the meaning 'every'.

\[(14)\] ikpà 'cup' ikpíkà 'every cup'
òwà 'house' ówòwà 'every house'
yèdè 'banana' èyèðéyèdè 'every banana'

Again, it is of course possible to describe these facts in a theory in which all features are contained in segment-size matrices. We could assume a feature analysis of the low, high, falling and rising tones as in (15) and write (16) to account for the tone changes. In prose, (16) says that a word-final vowel is deleted before a word-initial vowel, and if it has a different tone from the word-initial vowel of the next word, the tone on the word-initial vowel becomes rising if it was high and falling if it was low. While it correctly describes the situation, it presents the facts as arbitrary: we could just as easily write a rule with [−contour] rather than [+contour], but that rule would not describe any actual language data.

\[(15)\]

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
<th>Fall</th>
<th>Rise</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hightone</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Contour</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

\[(16)\]

\[
\begin{pmatrix}
\text{+syl} \\
\text{+hightone} \\
\text{−contour}
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\text{−syl} \\
\text{−hightone} \\
\text{−contour}
\end{pmatrix}
\rightarrow
\{+contour\}
\]

The autosegmental representation allows us to express the fact that it is only the vowel, not its tone, that is deleted. By reassociating the stranded tone to the next vowel, we accurately express the tonal adjustments. In (18b) a rise is produced, in (18c) a fall, while in (18a) no tonal change is predicted. Associating the same feature more than once to a structural element is meaningless, and the two identical tones are automatically reduced to one by a convention known as the Twin Sister Convention (see (17); Clements and Keyser 1983).

\[(17)\] Twin Sister Convention

\[
\begin{array}{c}
\text{\ x } \rightarrow \text{ } \text{\ x}
\end{array}
\]

\[
\begin{array}{c}
V \\
\text{T}_i \\
\text{T}_i \\
\text{T}_i
\end{array}
\rightarrow
\begin{array}{c}
V
\end{array}
\]

\[(18)\]

\[
\begin{array}{cccccccc}
\text{a} & \text{ikpa} & \text{ikpa} & \text{b} & \text{owa} & \text{owa} & \text{c} & \text{èye} & \text{èye} \\
\text{L} & \text{L} & \text{L} & \text{H} & \text{L} & \text{H} & \text{L} & \text{L} & \text{H}
\end{array}
\]

\[
\begin{array}{cccccccc}
\text{a} & \text{ikp} & \text{ikpa} & \text{ow} & \text{owa} & \text{èye} & \text{èye} \\
\text{L} & \text{L} & \text{L} & \text{H} & \text{L} & \text{L} & \text{H}
\end{array}
\]

\[
\text{∅ (by 17)}
\]
**Q96** Here are some examples of a word game in Bakwiri (Hombert 1986). Bakwiri has prenasalized stops: [mb], [nd] and [ŋ].

<table>
<thead>
<tr>
<th>Bakwiri</th>
<th>Word game</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a mòkò</td>
<td>kòmò</td>
<td>‘plantain’</td>
</tr>
<tr>
<td>b löwá</td>
<td>wáló</td>
<td>‘excrement’</td>
</tr>
<tr>
<td>c kóndí</td>
<td>ndikò</td>
<td>‘rice’</td>
</tr>
<tr>
<td>d ñòh</td>
<td>ñòñòñ</td>
<td>‘throat’</td>
</tr>
<tr>
<td>e ñzèè</td>
<td>zeñzèè</td>
<td>‘it is not’</td>
</tr>
<tr>
<td>f lúngá</td>
<td>ñgåålù</td>
<td>‘stomach’</td>
</tr>
<tr>
<td>g lióβá</td>
<td>βååljo</td>
<td>‘door’</td>
</tr>
</tbody>
</table>

1. What is the general rule for forming word-game words? Apply your rule to the hypothetical form [ndàk óó].
2. Bakwiri syllable structure is (C)V(V), where VV may be filled by one vowel, or by two different vowels. Give relevant examples.
3. Why isn’t the game form for ‘door’ *[βààlíó]*?
4. Why do you think there are glottal stops in the words for ‘it is not’ and ‘throat’?

**Q97** Kisanga has the syllable structure (C)V(V). In a children’s language game called Kinshingelo, sentences (a) and (b) are pronounced as (c) and (d), respectively (Coupez 1969).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>óbë múkë:tù t“á:já: kú múkólà</td>
<td>‘you, my friend, come with me to the river’</td>
</tr>
<tr>
<td>b</td>
<td>bà:nábáká:jí bá: mú kó::ñò b:tem“á: kúdímá</td>
<td>‘the Congolese women love to cultivate the land’</td>
</tr>
<tr>
<td>c</td>
<td>béó mútù:k“è ját“á: kú múlákó</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>bà:nábájíká bá: mú ñgò:kó b:ma“átë: kúmádí</td>
<td></td>
</tr>
</tbody>
</table>

Argue that this language game provides evidence for the existence in Kisanga of the word, the syllable, the mora and the tonal tier.

Tones may be systematically distributed over morphological categories along with morphemes that are also encoded in vowels and consonants. The term ‘grammatical tone’, one of the three functional categories distinguished in section 10.1, more strictly applies to **tonal morphemes** (Hyman 2016). Like intonational morphemes, tonal morphemes lack vowels and consonants in their phonological representation but, unlike intonational morphemes, have meanings that fit into the morphological and syntactic paradigms of the language, instead of expressing discoursal meanings.
or the prosodic phrasing structure. In Kalabari, which has a large number of tone melodies expressing morphological or syntactic categories, verbs can be made intransitive by imposing a LH melody on them (Harry and Hyman 2014). This is shown in (19), where ! indicates that the H in that syllable is downstepped (see section 10.7.1).

(19)  

<table>
<thead>
<tr>
<th>Transitive</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>a dimá ‘change’</td>
<td>dimá ‘be changed’</td>
</tr>
<tr>
<td>b sá’ki ‘begin’</td>
<td>sá’ki ‘begin’</td>
</tr>
<tr>
<td>c kikimá ‘cover’</td>
<td>kikimá ‘be covered’</td>
</tr>
<tr>
<td>d pákiri ‘answer’</td>
<td>pákiri ‘be answered’</td>
</tr>
<tr>
<td>e gbóló’má ‘mix up’</td>
<td>gbólomá ‘be mixed up’</td>
</tr>
<tr>
<td>f kán ‘demolish’</td>
<td>káán ‘be demolished’</td>
</tr>
</tbody>
</table>

Q98

1. What is the TBU in Kalabari?
2. What is the direction of association of the LH melody?

10.5 THE OBLIGATORY CONTOUR PRINCIPLE

The success of the autosegmental account of the word tone patterns in languages like Etung depends in part on the decision not to represent the pitch pattern of each syllable as a separate tone in the representation. Crucially, adjacent syllables with the same tone need not each be specified separately for tone, since the same tone can associate with more syllables through spreading. Considerations like these have led to the assumption that identical tones cannot be adjacent in the same word. In section 9.3.2, this restriction on representations was referred to as the Obligatory Contour Principle (OCP), formulated by Goldsmith (1976) with reference to Leben (1973) as (20). It forbids the occurrence of identical adjacent tones in the same morpheme.

(20) Obligatory Contour Principle:

$$\star V \quad V \quad T_i \quad T_i$$

Another argument for the OCP is that rules changing or deleting tones usually delete what on the surface appear to be sequences of identical tones. If we assume that the string of identical surface tones is in reality a single, multiply associated tone, a ready explanation is found for this phenomenon. Example (21), from Odden (1986), comes from Shona. The ‘associative prefix’ [né] ‘with’ lowers any consecutive
H-tones to low. Clearly, if these words have only one H, rule (22) can change that H into L, thus accounting for these data, as shown in (23).

(21) mbwá 'dog'
hóvé 'fish'
mbúndúdzí 'army worms'

(22) \( H \rightarrow L / H \) \\
    \( \ne \)

(23) ne-mbundudzi \( \rightarrow \) ne-mbundudzi

\[ \begin{array}{c}
    H \ \ \ \ H \ \ \ \ H \ \ \ \ L \\
\end{array} \]

In contrast, if each of the three syllables of ‘army worms’ were to have its own H, a more complicated analysis results. We must apply a rule deleting H after [né] provided the following tone is also H (which rule will apply more than once to the same word in a case like ‘army worms’), then change the remaining H-tone into L and, third, spread that L back to the preceding toneless syllables. Clearly, having a single H-tone in the representation makes life a lot simpler!

Q99 How does the derivation of ‘every cup’ proceed if, as argued by Leben (1978), Etsako obeys the OCP? Would there still be a need for the operation of the Twin Sister Convention?

In Optimality Theory, constraints are violable, and the OCP therefore cannot be a universally obeyed constraint. For one thing, languages like Etung frequently have exceptional patterns. For instance, words with surface melodies HHL and LLH do exist, like [ŋ̃gārē] ‘pepper’. Such words must be assumed either to have underlying melodies like HHL, contra the OCP, or to have a single H tone prelinked to two TBUs in the lexicon, contra the Association Convention. In the latter option, the surface representations of ‘pepper’ would look like (24) (Odden 1986). Representations that violate the OCP exist in many languages.

(24) ŋgāre

\[ \begin{array}{c}
    H \ \ \ \ L \\
\end{array} \]

10.6 ACCENT

Tones, whether lexical or postlexical, have a location and phonological content. The location of the tones in an Etung word is described in terms of the Association
Convention, while their phonological content is described in terms of a word melody. Another device for specifying location is accent. In section 1.2.2, we saw that Japanese words fall into two lexical classes, accented words and unaccented words. Accented words have a HL tone melody on one of their syllables, while unaccented words have no such melody. There are many generalizations about the presence or absence of the lexical HL tone melody, as in the case of compounds (Kuboizono 1993; Kuboizono et al. 1997). Certainly in informal descriptions, it will make sense to abstract away from the tones and assume some accent marking, so as to be able to refer to ‘accented word’, ‘accented syllable’, ‘accent deletion’, etc. It is not always obvious whether there is a theoretical advantage to an accentual analysis as opposed to a tonal analysis. When words are underlyingly marked for the presence of some specific tone or tones, such marking is effectively equivalent to the presence of lexical tone. Hyman’s (2001: 1368) definition of a tone language is suitably neutral between the two analyses: ‘a language with tone is one in which an indication of pitch enters into the lexical realisation of at least some morphemes’. Here, ‘an indication of pitch’ could refer either to an accent or to a tone.

The case for an accentual analysis is strengthened by the existence of more than one tone or tone melody in a language in combination with generalizations about tonal location. Barasana has two tone melodies, H and HL. These are associated left-to-right to moras, with spreading, much as in Zulu (Q94). Idiosyncratically, words begin the association at either the first or second mora (Gomez-Imbert and Kenstowicz 2000). A toneless word-initial mora will be pronounced with low pitch, equivalent to L. A tonal analysis would specify four tonal melodies, H, HL, LH and LHL, and assume a tonal association algorithm as for Etung. Representations in this analysis are shown in (25).

(25)  
\[
\begin{array}{cccc}
\text{a} & \text{b} & \text{c} & \text{d} \\
\text{jai} & \text{hee} & \text{wai} & \text{cai} \\
\text{H} & \text{H L} & \text{L H} & \text{L HL} \\
\end{array}
\]

‘jaguar’ ‘ancestral’ ‘fish’ ‘catfish’

An argument for an accentual analysis, effectively equivalent to that given by Gomez-Imbert and Kenstowicz (2000), can be based on the compound rule, which deletes tone on the second constituent, the head of the compound. To take (25a,b,c) as examples, [héè] ‘ancestral’ combines with [jáí] to form a compound [héè jáí] ‘shaman’, and [wáí] forms a compound with [kúmá] ‘glutinous stew’ to create [wáí kúmá] ‘glutinous fish stew’. This suggests that the second element in a compound loses its accent, such that only the tones on the first word are inserted. In an accentual analysis, lexical representations will need a specification of the accent (first or second mora), as well as of a tone melody, H or HL. The words in (25) now look like (26), where ‘*’ after a vowel indicates the accented mora.

(26)  
\[
\begin{array}{cccc}
\text{a} & \text{b} & \text{c} & \text{d} \\
\text{ja}^* & \text{he}^* & \text{wai}^* & \text{cai}^* \\
\text{H} & \text{HL} & \text{H} & \text{HL} \\
\end{array}
\]
Q100

1 Using the model of (25), give the tonal associations of Barasana ‘shaman’ and ‘glutinous fish stew’.

2 First and second person pronouns always have a H-melody, while third person pronouns have a HL melody. When prefixed to a noun, their tone melody copies onto the noun, whose original tone melody is deleted, as in the examples below. Use this rule to argue for an accentual analysis.

   a  mání  ‘1PL’ míni  ‘pet’  –  mání míni  ‘our pet’
   b  ínà  ‘3PL’ míni  ‘pet’  –  ínà míni  ‘their pet’
   c  jìi  ‘1SG’ míni  ‘pet’  –  jìi míni  ‘my pet’
   d  mání  ‘1PL’ wìhìbò  ‘tray’  –  mání wìhìbò  ‘our tray’
   e  ínà  ‘3PL’ bábářì  ‘friends’  –  ínà bábářì  ‘their friends’
   f  jìi  ‘1SG’ wìhìbò  ‘tray’  –  jìi wìhìbò  ‘my tray’

All English words are in principle accentable, as indeed they must be, since an utterance contains minimally one accented word, whose main stressed syllable will host an intonational tone or tone melody. The analytical separation of the location of the tone and its identity is the only option in a case like English. First, all tone assignments are postlexical, which implies that any lexical specifications of tones are ruled out. If we were nevertheless to insist on specifying tones in the lexicon, our efforts would be frustrated by the fact that there are many tone melodies that can in principle be inserted in an accented syllable, depending on their discoursal meaning. Also, there are many generalizations that can be made about the assignment or deletion of pitch accents. It would be bizarrely uneconomical to list the tones that could occur in an accented syllable as being absent from it, instead of just saying that the syllable is not accented. One deaccentuation rule of English is the compound rule, which is just like the Barasana rule: delete the accent on the second constituent (27). Another is a deaccenting rule for post-focal words, whose effect is seen on the lack of accent on pills in (28b), where the focus is don’t need. The reason for a deaccentuation can on occasion be ambiguous. The title of Walter Lang’s 1954 movie There’s NO Business Like SHOW Business sounds as if the second mention of business is deaccented because of the previous occurrence of that word in the sentence, but it would of course have been deaccented anyway, as it is in There’s NOTHING like SHOW business, because it is the second constituent in a compound.

(27)  a  BUSiness, MODel  –  BUSiness model
     b  DRAma, QUEEN  –  DRAma queen
(28)  A: Did you get your PILLS?
      B: I don’t NEED those pills anymore.
The description of the pronunciation of tones follows the target-and-interpolation model outlined in section 8.6.2. A tone’s pronunciation is thus a target with a ‘vertical’ and a ‘horizontal’ specification. The ‘vertical’ specification is some value on the f0 scale, known as its ‘scaling’ (Pierrehumbert 1980: 47). While Pierrehumbert (1980) concentrated on the scaling of tonal targets, most of the subsequent research has been devoted to the ‘horizontal’ dimension, the synchronization of the tonal target with the segmental string. The phonetic implementation of tones, just as of vowels and consonants, is both context-sensitive and language-specific, as explained in section 8.6.1. To illustrate, consider the phonetic implementation of the Mandarin lexical tone patterns in (1), which will have tonal representations much as in (29). The bracketed H in the Low word tone in (1c) only shows up when another L follows and at the end of a phrase (see Duanmu 2000: 221). The end of the syllable provides the approximate location of the last tone’s target (Xu 2006). A sequence of a Low tone (L), a High tone (H) and a Falling tone (HL), therefore, will be pronounced as shown by the solid line in panel (a) of Figure 10.1, where the bullets represent the targets of the tones. This means that the High tone comes out phonetically as a rising pitch movement through the second syllable. However, this pitch contour is distinct from the pitch contour for the Rising tone (LH) in the same context, whose two targets are indicated by the open circles. The rising movement for LH, which is defined by the interpolation between a low target and a high target in the same syllable, is located late in the second syllable, instead of covering the entire syllable. In panel (b), the averaged measured contours are shown for the second syllable in emphatic speech. In this speech style, the high target of the rising tone may in fact be located just over in the next syllable. This ‘peak delay’ is a well-known feature of tonal realizations in languages more generally.

Figure 10.1
Theoretical Mandarin Chinese pitch contours of High tone and Rising tone between Low tone and Falling tone (panel a), and actual pitch contours of High tone and Rising tone after Low tone and before Falling tone, averaged over 18 tokens (panel b; adapted from Chen and Gussenhoven 2008).
The f0 tracks in Figure 10.1 make it clear that we cannot think of the phonetic realization of tones as sequences of high and low pitches. Just as sequences of vowels will be pronounced with transitions between their acoustic targets, so will the targets of tones be separated by transitional phases, the interpolated stretches of the f0 contour.

10.7.1 The vertical dimension: scaling

Factors affecting the scaling of tonal targets which are likely to play a role in many languages are upstep, downstep, H-raising, declination, final lowering and undershoot.

Upstep

When the second of two adjacent H-tones is pronounced at a higher pitch than the first, it is upstepped. Upstep is rare for lexical tones, but was reported for Acatlán Mixtec by Pike and Wstrand (1974), where the implementation rule is iterative and thus upsteps a H also relative to a preceding upstepped H. Non-iterative upstep is common in intonation systems. Intonational tones are independent morphemes. A distinction is made between tones that are associated with the accented syllables of some words, and tones that are located at the edges of larger domains. The former are word melodies, or pitch accents (Pierrehumbert 1980), while the latter are known as boundary tones. Their meanings have to do with whether you are giving the listener some information or are asking for some and how the listener is supposed to relate the information with what she or he already knows (e.g. Pierrehumbert and Hirschberg 1990; Bartels 1999), while boundary tones in addition indicate prosodic phrasing. Like word melodies, pitch accents may consist of one or more tones, but unlike the word melodies of Etung and Barasana, typically only one of them associates with the accented syllable. Because the associating tone may be any of the tones in a multi-tone pitch accent, it needs to be marked for its preferential association, which is done with the help of a following ‘*’2 In Western European languages, boundary tones occur at the edges of the intonational phrase, with a percent sign indicating the edge, %L and H% in (30). The example shows the application of upstep to H% after a preceding H* (Pierrehumbert 1980).
Downstep

Downstep amounts to a lowering of a H-tone relative to a preceding H-tone, in some specific context. There are broadly two phonological contexts in which it occurs. One requires a L to occur between the triggering H and the H that undergoes downstep. The other requires the trigger to be adjacent – on the tonal tier, i.e. disregarding the segmental tier! – to the H undergoing the downstep (Odden 1986). An example of the first occurs in Kuki-Thadaw, which has three word melodies, HL, H and L, and uses the syllable as the TBU (Hyman 2007). HL surfaces only as a fall in phrase-final position. When occurring before another tone, a level high tone is realized, which is accounted for by delinking L. A tone that has no association with a TBU is said to be floating. In (31a), two words with HL are followed by words with L, creating two valleys, whereby the second, third and fourth peaks are downstepped relative to the preceding one. Observe that in (31a), the floating L-tone may be assumed to be deleted. Its presence has no effect on the realization of any other tone. By contrast, in (31b), the presence of a floating L is revealed by the downstep it triggers on the next H. The second and third words have a downstepped H, without a preceding valley, and the presence of the floating L-tones is revealed only through the lowered pitch of the next H. The effect of the floating L is confirmed by the form in (31c). As in many languages, a H spreads right in Kuki-Thadaw, in this case before L. H-spreading applied in (31c) so as to disassociate the lexical L, a second source of a floating L. As predicted, the following H is downstepped.

Iterative downstep, as in (31a,b), is in principle unbounded. A classic study by Liberman and Pierrehumbert (1984) shows how the progressively downstepped H-targets can be modelled with the help of a fraction of the portion of the f0 above a reference line. With a reference line of 100 Hz, a preceding target of 150 Hz and a downstep factor of 0.7, this will place the next H-target at 100 + (150 – 100) × 0.7, or 135 Hz.

Downstep of H after H occurs in Cilungu (Bickmore 2007). Toneless [ful] ‘to wash’ in (32a) is preceded by the class prefix [ku] and the subject [tú] ‘1pr,’ the only toned morpheme in (32a). Its H spreads to the toneless syllables till the end of the verb, leaving [a] toneless, a morpheme known as the Final Vowel, which is generally present on verbal forms. In (32b), a H-toned verb [sopolol] ‘to untie’ is used in the same context. Now the H of [tú] spreads until the H of the verb, which itself spreads till it gets to the Final Vowel.
H-raising

H-raising is common in African languages and has notably been investigated for Yoruba, which contrasts H, M and L in the surface representation. Before L, but not before M, H is raised. The effect can be so strong as to cancel out the effect of downstep, meaning that the pitch of the second H-tone in a HLHL sequence may barely be lower than the first (Connell and Ladd 1990; Laniran and Clements 2003). H-raising is also sometimes referred to as 'upstep'.

Declination

Declination, shown in (33a), applies over larger constituents like the intonational phrase and amounts to a gradual lowering of tone targets across the domain, an effect which is weaker than that of downstep, if that exists in the language. Thus, even though a sentence may contain exclusively H-tones or exclusively L-tones, a descending trend is typically observed. In questions, declination may be suspended (Grønnum 1992).

Final lowering

The last 400 ms or so of a declarative utterance are lowered over and above the effects of declination and downstep. The extra lowering of plans in (33a) may serve as a finality cue, indicating an opportunity for a turn switch to another speaker. In (33b), a downstepping contour is shown in which the targets of the final H* and L% lie below the rates of descent established for the preceding high and low targets in the sentence (Liberman and Pierréhumbert 1984). Without downstep, the H*-targets may still show a descending pattern due to declination, an effect that is much smaller and possibly less systematic, as illustrated in (33a).

Final lowering may resemble downstep when it affects a final H-tone in Dinka. The language has lexical L, H and HL and employs the syllable as the TBU, as in [tʃːk] ‘woman’, [bɔŋ] ‘chief’, [tʊŋ] ‘horns’ (Andersen 1987). Dinka downsteps H after a L relative to an earlier H. In (34), the word [lɔːk] (Dinka has a three-way vowel quantity contrast) is pronounced with downstepped H because of the preceding
HL, however, the phrase-final syllable also has a lowered tone which, according to Andersen (1987), is like that of a downstepped H. Since the syllable is phrase-final, a final L% may be held responsible for the lowering. This boundary tone may be assumed to associate with the last syllable, so that, together, H and L% are realized as a lowered H, rather than as falling pitch.

![Diagram](34)

`ca megl n j n e tr:n`

H L H H H L%

PERF; 2SG child wash; PARTICIPLE when here

“When did you wash the child here?”

Undershoot

When time is at a premium, as will happen when a number of different targets need to be reached in a short time, the more extreme targets may be undershot. This will happen for the end point of the raising and lowering of the tongue in a sequence like /aia/, which may be pronounced [aea] or [aea] as a result. Similarly, a HLH sequence may come out as [HMH]. An extreme case of this occurs in Zeeland Dutch, where H*L H% on a final syllable in an intonational phrase may be realized as a sequence of two rises with a brief level stretch between them (Hanssen 2017).

10.7.2 The horizontal dimension: alignment

Associated tones are pronounced close to their TBU. There is a general tendency for H-tones to be delayed, causing their target to occur late in the syllable or early in the next syllable. Delayed H-targets may cause a following L to ‘fall off’ the segmental utterance, causing truncation of falls. Grabe (1998) shows how a long-vowelled German syllable *schief ‘slanted’ will show a fall to low pitch in the same communicative context in which short-vowelled *Schiff ‘ship’ shows high pitch or a fall from high to mid. The tendency to delay H is reflected in right-spreading of tones in African tone languages in particular. For instance, in Yoruba L-H is restructured as L-LH and H-L as H-HL (Hyman 2007).

Floating tones may not in fact have a target. For one thing, languages may routinely delete them. When they are retained, their presence may only be felt in the target of a neighbouring tone, as discussed above under ‘Downstep’. Floating intonational tones are typically retained and pronounced. In Figure 10.2, the pitch contour and speech waveform are given of You may MINimize the email window by clicking on that DASH there. There are two occurrences of the pitch accent H* L; the sentence begins with %L and ends with H%, as shown in (35). Of these six tones, two are associated, while four are floating. In English, only one tone can associate with the TBU, the accented syllable, the ‘starred’ tone, H* in this case. The timing of associated tone H* is relatable to a specific syllable, but there is no invariant syllable with which the following L is timed. In a final syllable in the intonational phrase (IP,
see chapter 12), it occurs on the same syllable as H*, but in (36) it may only occur on can, two syllables after the accented syllable. It wouldn’t make much of a difference if L was pronounced during you or even see, but a similar displacement of H* would be very noticeable and would create the impression of a different intonation pattern. That is, the timing of the pitch accent is determined by the anchoring of the starred tone to the rhyme of the accented syllable. Unsurprisingly, %L and H% are realized at the beginning and end of the IP. Boundary tones illustrate par excellence how tones, like all phonological elements, are aligned somewhere in the linguistic structure (see also section 11.3.2; McCarthy and Prince 1993). An alignment statement specifies the edges of two constituents which should temporally coincide. In English, the right edge of H% aligns with the right edge of the IP. Trailing tones of non-final pitch accents, like L of H*L on minimize in (35), align rightmost, being pronounced just before the target of H*. By contrast, trailing tones of final pitch accents are pronounced leftmost, here causing a steeper fall in H*L. If we expand the IP with unaccented words, as in (36), the fall from DASH would still be steep and you can see or can see will more or less have low-level pitch throughout.
Q101 Describe the f0 pattern that will occur on John if the sentence Has JOHN? is pronounced with the intonation contour %L H*L H%. Give a representation with tones and targets in the way shown in (35) and (36).

Q102 In Kipare, an utterance-final H is downstepped, or in our terms finally lowered, as shown in (1a) (Compare (1b); Odden 1986).

(1)  
a ipáŋgá  ‘machete’
b ipáŋgá lédi  ‘good machete’

Assume that the end of the utterance has a boundary L%, which causes the downstep on the final H. When L-tones precede the final H, a frequent, though not obligatory, pronunciation is one in which the final (downstepped) H spreads leftward to all the L-toned TBUs. (2a) shows the pronunciation without the spreading, (2b) the pronunciation with.

(2)  
a nìfinikirè màyémbè màë’dá  ‘I covered long hoes’
b nìfï’níkiré màyémbè màëdá

Although it is optional, the spreading of the final H shown in (2b) is always to all TBUs. That is, it is not possible to pronounce the sentence with spreading of H to only some of the words that are pronounced with low pitch in (2a). What does this suggest about the representation of the L-toned words before H spreads?

The term ‘alignment’ has been used for the phonological specification of a tone’s location (Pierrehumbert 1980; Gussenhoven 2000) as well as for the detailed phonetic specification of a tone’s target (‘phonetic alignment’; Arvaniti et al. 2000). Languages differ in the typical location of both associated and floating tones. One factor that has been shown to cause retraction of H-targets is the proximity of an upcoming H-target or IP boundary. The need for sufficient space to carry out a pitch movement may therefore counteract the tendency to allow H-targets to drift rightwards.

10.8 PRIVATIVE TONE

Tone systems often lend themselves to a privative analysis. There are two ways in which a contrast can be privative. Most frequently, the tone contrast exists on the
surface, but, underlyingly, there may be no need to specify one of the tones involved in the contrast. For instance, in a language with H and L, only H might ever be referred to by phonological generalizations. When all the constraints of the grammar have been taken into account, any remaining toneless syllables are provided with L tone.

The Somali word has maximally one H-tone, always in the final two syllables. TBUs without a H-tone are pronounced at low or mid pitch, which we may assume is due to a default L added at a final stage to free moras, the TBUs of Somali. In the grammar, there are many rules positioning the H-tone, but no rules other than the default rule ever refer to L.

### Q103

Somali uses tone to encode a gender contrast (Hyman 2009). Masculine nouns have falling tone on final long syllables, low tone on final short syllables and, depending on the vowel quantity in the last two syllables, high, low or rising tone on the penult, while feminine nouns have high tone on final short vowels, rising tone on final long vowels and low tone on the penult.

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>'boy'</td>
<td>'girl'</td>
</tr>
<tr>
<td>'young he-goat'</td>
<td>'young she-goat'</td>
</tr>
<tr>
<td>'stupid man'</td>
<td>'stupid woman'</td>
</tr>
<tr>
<td>'young male camel'</td>
<td>'young female camel'</td>
</tr>
<tr>
<td>'colt'</td>
<td>'filly'</td>
</tr>
<tr>
<td>'liver masc'</td>
<td>'garden fem'</td>
</tr>
</tbody>
</table>

Can you improve on the above description of the two tone patterns?

The privative opposition may also persist on the surface. In the Central Franconian dialects of German and Dutch, about half the words have a lexical tone, other words being toneless. The languages have a generous supply of intonational tones, to which one or more lexical tones are added depending on the number of words with lexical tone in the sentence. If no such words are included, the sentence just has intonational tones. In Hasselt Limburgish, the lexical tone is L, while LH* is the only pitch accent. In (37a), only intonational tones occur, while in (37b) the word for ‘socks’ has a lexical L-tone in the stressed syllable. Because the TBU is now occupied, the pitch accent, which is aligned to the right of L, has no TBU to associate with. Phonetically, the pitch peak due to H in (37b) is later than that in (37a) (Peters 2008).

(37) a ix mut nɔx an paar kiaskas hebo %L L'H L% ‘I still need a few cheeses’ b ix mut nɔx an paar kiaskas hebo L% L'L'H L% ‘I still need a few socks’
10.9 NOT BY f₀ ALONE

Vocal folds are not just employed for creating variations in their rate of vibration. Voice quality differences are effected by different modes of vibration, while voicing contrasts in consonants provide yet another use in the phonologies of languages. In sections 10.9.1 and 10.9.2, we discuss the connections between these two aspects and tone.

10.9.1 Voice quality

Modal voice may phonologically contrast with either breathy voice or creaky voice or both (see section 2.2.3). Breathy phonation is associated with lower rates of vocal fold vibration, since the airflow through the glottis is to a large extent usurped by the need to create friction between the vocal folds. A depressing effect of breathy voice is seen in Itunyos Trique (DiCanio 2012), where breathy voice occurs as a result of [i] in the coda. Compared to the f₀ patterns in open syllables, the f₀ patterns of the four level tones descend in syllables closed by [i], while the mid falling contour descends more steeply. Interestingly, in addition to these five tones, open syllables have two falling contours, whereas syllables closed by [i] have two rising contours. Apparently, having seven contrastive falling contours goes beyond what is ergonomically tolerable for the language.

Creaky voice presents a more complex picture. Often, it is associated with low tones. In Standard Mandarin, Tone 3, the low or dipping tone, routinely has creak, which helps to make it distinct from the rising tone, Tone 2. Creak as an enhancement of low pitch is a frequent feature in languages, a connection which must stem from the ease with which irregular, creaky phonation can be achieved with lower vibration rates. At the same time, tenseness of the vocal folds as used in ‘constricted’ or ‘tight’ phonation will lead to higher vibration rates. Kingston (2005) presents the development of laryngeal segments into high and low tones in different dialect groups of Athabaskan as resulting from these opposite effects. Co-occurrences of phonation types and lexical tones are common in varieties of Vietnamese (Michaud 2004; Brunelle 2009). When creaky voice derives from the occurrence of [ʔ] in the coda, a shortening of the sonorant portion of the rhyme frequently occurs. Since pitch shapes are less distinguishable on short vowels, a reduction in the number of tone contrasts typically results. For instance, the seven-tone contrast in non-glottal rhymes in Itunyos Trique is reduced to a three-way contrast in glottal rhymes.

10.9.2 f₀ perturbations and tone distribution

Consonants can ‘perturb’ the smooth phonation which is characteristic of vowels and sonorant consonants, with their free flow of air from the glottis onwards. The oral constriction of voiced obstruents like [b] or [v] will cause air pressure to
build up in the oral and pharyngeal cavities, as a result of which f0 rates may be depressed, because the higher air pressure above the larynx will inhibit the flow of air through the glottis. The effect is often visible as a dip in the f0 track. The depressing effect may carry over to following vowels, whose f0 will be lower in the first 20 ms or so than after sonorant consonants like [m]. Voiced implosive consonants like [ɓ ɗ ɠ] may have a stronger such effect. Vocal fold vibration here relies on the lowering of the larynx, which will result in a simultaneous subglottal air pressure increase and supralaryngeal air pressure decrease. This pressure difference is used to get the vocal folds to vibrate, but the effect is short-lived and weak and leads to a low vibration frequency. Voiceless obstruents, by contrast, are likely to have an f0 raising effect due to the tensing of the vocal folds which is created in order to keep the glottis open. Tone systems may reflect these effects in segment dependent distributions. Languages in Africa frequently have low tones after implosives, and varieties of Wu Chinese tend to have the full suite of tone contrasts only after sonorant consonants. The Shanghai variety has three tones which appear on sonorant rhymes, HL, MH and LM, but they contrast only after sonorant onsets. After voiceless obstruents, only the higher two tones occur, HL and MH, while voiced obstruents only tolerate LM, the lowest tone. In rhymes closed by a glottal stop, which have very short sonorant portions in the rhyme, there is a two-way contrast between H and L after sonorants, while there is no contrast after obstruents, with H occurring after voiceless ones and L after voiced ones (Chen and Gussenhoven 2015). The Yuhuan variety has more tone contrasts in sonorant rhymes, as shown in (38), but shows a similar distribution. That is, after voiceless obstruents, H, HL and LH appear, while L, ML and LH appear after /b d g dz v z fi/, the voiced obstruent onsets (Gussenhoven et al. 2016).

(38) a miɔ b miɔ c miɔ d miɔ e miɔ
H L HL ML LH
‘glance’ ‘temple’ ‘second’ ‘depict’ ‘seedling’ (unit of time)

Q104 Yuhuan Wu syllables can begin with [j] and [w], optionally preceded by a consonant. These glides are either to be analysed as onset consonants or as initial vowels in the rhyme, so that the segmental structure of the LH word for ‘rice seedling’, phonetically [jà], is either /jà/ or /ià/. Syllables of this type have a three-way tone contrast.

1 What would you consider the correct analysis of the glides?
2 Which three tones do you expect to occur in these syllables?
3 And how many tones can appear on a syllable like /i/?
Q105 Giryama has a H vs Ø contrast on words, with default L-insertion in toneless syllables. An underlying H-tone is pronounced on the penultimate syllable of the intonational phrase, regardless of the position of the morpheme it comes with. In (1a) no H-tones occur, while in (1b) the 3sg prefix /á/ occurs in initial position. However, its H-tone associates with the penult, where a somewhat delayed pronunciation causes that syllable to have rising pitch (Volk 2011).

1 Explain the difference between association and (phonological) alignment on the basis of (1b), and be explicit on whether H spreads rightwards from its sponsoring morpheme.

2 What might be the explanation of the association of H with the penult rather than the final syllable in (1a,b)?

3 When two H-tones occur, the first associates with the first syllable of the verb, as in (2a), which contains a verb with an underlying H. What might be the reason that the first H-tone in (2b) associates with the syllable preceding the (H-toned) verb?

4 Does the constraint on the association of the first H with the verb equally apply to the final H in the IP?

(1) a nì nà màrigizi:kà  
1st pres come-to-an-end  
‘I am coming to an end’

b à nà màrigizi:kà

‘She or he is coming to an end’

(2) a à nà kúlihiri:kà  
‘She or he is trustworthy’

b à ná bàmbahú:là

‘She or he is separating’

10.10 CONCLUSION

This chapter has dealt with a conspicuous phonological type of segment, tones. We have seen how tone in African languages gave rise to the autosegmental model of phonological representation and the main constraints that govern those representations, like the No Crossing Constraint, the Twin Sister Convention, the No Crowding Constraint and the OCP, as well as association algorithms like the Association Convention. Tones can be distinguished on the basis of their function: they may be lexical, grammatical and intonational. Phonologically, they may be associated or floating, but always located somewhere, depending on their phonological alignment. We have seen that descriptions of the distribution of tones may call on concepts like tone melody and accent. The phonetic implementation of tones was discussed in terms of their scaling, the placement in the f0 range, and their synchronization with the segmental structure, the phonetic alignment. For each dimension, a number of widely attested factors were identified that
systematically affect the realization of tone targets. Finally, we have drawn attention to the privative nature of tone oppositions and the way tone systems may be shaped by the perturbations of f0 caused by consonantal articulations. The next chapter tackles the third aspect of word prosodic structure, stress.

NOTES

1 See http://www.oto-manguean.surrey.ac.uk/ for a survey of linguistic properties of Oto-manguean languages.

2 Additionally, multi-tone pitch accents have a ‘+’ between their tones in Pierrehumbert (1980) and a great deal of later work, as in L^*+H or L+H^*, a convention that is not adopted below.
11.1 INTRODUCTION

Besides lexical tone and quantity, words may have stress. This is a phonological version of an articulatory tendency to pronounce syllables with an alternating strong-weak pattern. For some reason, it is hard to pronounce a sequence of syllables with equal care for each of them. For instance, when reciting the alphabet in a language with monosyllabic words for the letters, the tendency may reveal itself in a strong-weak pairing of the letter names: Á B Ć D Ê F Ĝ. In many languages, this tendency is reflected in the phonological structure of words, known as a ‘metrical structure’, a ‘foot structure’ or simply ‘stress’.

There are phonological as well as phonetic diagnostics for stress. Hyman (2006) identifies three phonological properties that characterize word prominence as ‘stress’:

1. The word prominence is obligatory. Languages with word stress have stress on all of their words (except functional morphemes, which may be incorporated with other words; see 12.6.1 on clitics). As an obligatory property of words, stress differs from tone or bimorcity.

2. Stress is a property of a syllable. In particular, there is no such thing as a stressed mora as distinct from a syllable.

3. Stress is culminative, meaning that one syllable in the word has the greatest prominence.

Phonetically, stress has been defined as increased articulatory precision as shown by an increased segmental duration, more peripheral vowel qualities and greater intensity, in particular for the higher spectral frequencies (Fry 1955, 1958; Beckman 1986; Sluijter and van Heuven 1996). The employment of f0 features is seen as a reflection of the presence of tones on stressed syllables.

In section 11.2, we consider how stress manifests itself in languages, keeping the above distinction between phonological and phonetic criteria in mind. We will discuss tone systems that resemble stress systems in a distributional sense, making the point that such systems have sometimes been described as stress systems, as well as languages that have neither tone nor stress. Tones, whether or not analysed as pitch accents, are more frequently deleted postlexically than stress and are as a result more likely to figure in morphosyntactic generalizations. Section 11.3 will make the point that stress systems always have a default location, meaning that there are no clear examples of languages which distribute stress entirely unpredictably. A language need not have a single default location, like the penultimate syllable, but may
vary the default as a function of the segmental make-up of syllables that potentially have stress, a phenomenon known as quantity-sensitivity or weight-sensitivity. Importantly, we will see how the formulation of generalizations about the location of stress has been an exciting area of research in the past decades. We identify and define the main parameters that figure in these generalizations on the basis of data from a large number of languages. Finally, section 11.4 shows how Optimality Theory has been applied to the problem of deriving default stress distributions.

### 11.2 Manifestations of Stress

Syllables with stress may distinguish themselves from other syllables in a number of ways. In many languages, stressed syllables allow greater phonological complexity, another case of phonological privilege (‘positional faithfulness’), besides the position of the syllable onset (section 3.2; Beckman 1988). First, they are the syllables to which tones associate. These tones or combinations of tones are also known as ‘pitch accents’, a somewhat overused term which should be used with care to avoid ambiguities. As illustrated in section 10.7, English pitch accents make up the intonation of sentences together with boundary tones (Pierrehumbert 1980). A common pitch accent of English is a rising-falling peak, which may be analysed as a falling melody (H*L; e.g. Leben 1976). An utterance in English must minimally have one pitch accent, and a citation pronunciation of Edinburgh, which has the main stress on the antepenultimate syllable, [ˈɛnθəbrə], will therefore have a pitch accent (here H*L) associated with the first syllable, as seen in (1a). Because in Alaska [əˈlæska] the second syllable has the main stress, this is where the pitch accent goes, as in (1b). Aberdeen [ˌæbəˈdiːn] has two stressed syllables, the first being a ‘secondary stress’. While the main stress must have a pitch accent when the word is said by itself, a secondary stress before the main stress may also have a pitch accent, creating two pitch accents in this word, as shown in (1c). About half of the stressed syllables in an English text will be accented, i.e. have a pitch accent.

![Diagram of pitch accents]

Just as only stressed syllables may have a pitch accent, stressed syllables may be segmentally privileged. They uniquely have bimoraic syllable rhymes in Swedish, where bimoraic quantity for vowels is mutually exclusive with geminate consonants (e.g. */at/, /æt/, /at/, */aːt/; Riad 2014). Catalan has seven vowels in stressed syllables but only three in unstressed ones (Nadeu 2016), and most varieties of English have large numbers of vowels, but only a small number of vowels can appear in unstressed syllables (Bolinger 1986: 347; cf. Gussenhoven 2011). When stressed syllables are signalled by a pitch accent, but do not differ from unstressed ones in segmental complexity, they may still differ in duration or vowel quality. For instance,
Spanish unstressed syllables, which don’t obviously distinguish themselves from stressed syllables in their segmental composition (Berg 1998: 112), are shorter than stressed syllables (Ortega-Llebaria and Prieto 2011), while the same will be true for many other languages without enhanced segmental complexity in stressed syllables. In all of these cases, the stressed syllables are said to have phonetic stress. It means that besides f0 features which are attributable to the presence of stress, there are observable durational and spectral properties, whether or not these are due to phonetic realizations of segmentally similar syllables, as in Spanish, or to systematic segmental differences between stressed and unstressed syllables, as in English. In fact, the distinction between ‘pure’ phonetic stress and segmentally induced phonetic stress may be hard to draw, because it may depend on the analysis. For instance, what one phonologist may regard as the stressed and unstressed allophones of the same vowel, another phonologist may take to be two different vowels.

Q106 Up until the 14th edition of Daniel Jones’s English Pronouncing Dictionary (Jones 1977), the words forward and foreword were given as [ˈfɔːr.wəd] and [ˈfɔːr.wɔːd] (ignoring other differences), while in the 14th edition, which was fully revised by his successor A.C. Gimson, they have been given as [ˈfɔːr.wəd] and [ˈfɔːr.wɔːd]. How can this difference be interpreted phonologically?

Stressed syllables may equally be the privileged locations of lexical tones. In the Cologne dialect of German, the main word stress is the location of a tone contrast (Gussenhoven and Peters 2004), while the same is true for word-final stressed syllables in Ma’ya, which has a three-way contrast there (Remijksen 2007), and Swedish non-final syllables, which like Cologne German have a two-way tone contrast (Riad 2014).

11.2.1 Absence of phonetic stress

What if a language has no phonetic stress, i.e. no phonetic properties apart from the f0 effects of the pitch accent, but distributes f0 features in a way that is reminiscent of the way stress is distributed across words? In Persian, for instance, every major-class word as pronounced in isolation has high pitch on the last syllable, which is evidently due to a H-tone, but because it has no phonetic stress (Abolhasanizadeh et al. 2012), any motivation for saying it has ‘stress’ lies purely in the distributional property of the tone: every word has one, at least when said in isolation. Beckman (1986) avoided any possible confusion by distinguishing syllables with phonetic stress as having ‘stress accent’ and syllables without as having ‘non-stress accent’, respectively. Under ‘stress accent’, she explicitly included cases in which stressed syllables differ from unstressed ones in phonological complexity, like English, while Japanese was her example of a language with ‘non-stress accent’. Still, instead of ‘non-stress accent’, many linguists will use the term ‘pitch accent’ here. In this usage of that term, it is important to keep in mind that reference is made to a language
without phonetic stress, and that ‘pitch accent’ here means ‘exclusively tone’. Recall that English, a language with obvious stress, also has pitch accents. Thus, the presence of a pitch accent does not in general exclude the presence of phonetic stress.

When tones are deleted from a syllable in a language without phonetic stress, that syllable is phonologically identical to a syllable without tone at any representational level. To illustrate the homophony of ‘deaccented’ syllables and syllables that never had a tone, we turn to Persian again. The generalization that the last syllable of every major-class word has a H-tone is true regardless of the attachment on the right of function words, like auxiliary verbs and possessives. For instance, [tə拜ʃ], the proper name Tabesh, is distinct from [tə拜ʃ] ‘my swing’, where [əʃ] is the 1sg possessive. The language has a number of rules deleting tones (‘pitch accents’). One of these applies in object clauses of factive verbs (Rahmani et al. forthcoming). In this context, the two above structures are homophonous, causing (2) to be ambiguous.

(2) mehdi midune ke un tə拜ʃe
   a Mehdi knows that that+DEM Tabesh-is
     ‘Mehdi knows that that is Tabesh’
   b Mehdi knows that that+DEM swing-my-is
     ‘Mehdi knows that that is my swing’


Q107 Nubi has no phonetic stress. Every word is lexically marked for the location of a H-tone appearing on the antepenult, penult or final syllable, referred to as the ‘accent’. The language has two rules that manipulate accents. First, a gerund with a direct object is deaccented, as shown by the contrast in (1) and (2).

(1) Ána gú rúa wedí buká
    I will give books ‘I will give books’
(2) Wedí buká séme giving books good ‘Giving books is good’

Second, an adjectival verb before a noun loses its accent in favour of an accent on its first syllable, as illustrated in (3) and (4).

(3) tabán ‘to taunt’
(4) tában tóru ‘to taunt a bull’

1 What are the Nubi translations of ‘I will taunt a bull’, ‘I will taunt’ and ‘Taunting a bull is good’?
2 Give two arguments for characterizing the Nubi accent as tone, not stress.
11.2.2 Phonetic stress as an obligatory feature

Even though Hyman (2006) defines ‘stress’ in terms of distributional properties and is thus neutral on how it is realized, the languages with ‘obligatory’ stress he discusses would all appear to have phonetic stress. Apparently, phonetic stress is undeletable, i.e. obligatory, and invariably shows up in surface representations. English \textit{foreword} [\textipa{fɔː(r)wɔː(r)d}] and \textit{forward} [\textipa{fɔː(r)wɔ(r)d}] are distinct in any phonological context, even if no tones (pitch accents) appear on their stressed syllables. In contrast, tones tend to be manipulable and are thus frequently targets of phonological rules, including deletion rules.

Languages vary in the way they exploit tone on stressed syllables for expressing linguistic contrasts. Untypically, from an Indo-European point of view, every Egyptian Arabic word in the sentence has a pitch accent on its stressed syllable (Hellmuth 2007). By not being distinctively present or absent, the only function it has is to signal the presence of a word. At the other extreme, English has many grammatical distinctions which are expressed through differences in accentuation (cf. section 10.6). For instance, \textit{The STORKS are RINGED} counts as a general rule of practice, while \textit{The STORKS are being ringed} might be used to report on the actual process of people providing some storks with tracking rings.

11.2.3 Lexical vs postlexical prosody

Discussions of word prosodic structures should strictly speaking be confined to those that are included in lexical representations. For tone, this is typically made explicit. Definitions of the notion ‘tone language’ invariably include some such provision, and postlexical tones are separated off as intonational tones. However, postlexical tones (whether or not they are called ‘pitch accents’) have been referred to as ‘stress’ when their distribution resembles that of stress. For instance, French words are sometimes said to have word-final stress as a result of the fact that the language has pitch accents on the initial and final syllables of phonological phrases, and a citation pronunciation of a word will therefore have a pitch accent on its final syllable. ‘Real’ stress is postlexical in languages that assign feet to phonological words, without any exceptions. For instance, Bengali has stress on the first syllable of every major-class word and rhythmic stress on every odd syllable in the word thereafter (Chatterji 1921: 19). The only clearly attested manifestation is the presence of a pitch accent on the first syllable of the first such word in the phonological phrase (Hayes and Lahiri 1991; Kahn 2010), but the report of binary rhythmic stress by Chatterji (1921) suggests that the language has postlexical stress with a phrasally determined postlexical pitch accent, the latter aspect being comparable to French. Parallel to lexical tone, stress is lexical if its location is sensitive to morphological operations or if segmentally comparable underived words have different stress locations (see section 11.3.4). Postlexical quantity occurs in Italian (by the side of lexical quantity due to its intervocalic geminate-singleton contrast, illustrated by [\textipa{di.ta}] ‘fingers’ vs [\textipa{di.tə}] ‘business company’). If the last word in an intonation phrase has penultimate stress on an open syllable, the vowel is lengthened, as in (\textit{Vivo a}) \textit{R[ɔː]ma} ‘I live in Rome’. The lexical status of prosodic phonological
features may be detectable in the results of a short-term memory task involving prosodic minimal pairs like [númi – numí], developed by Emmanuel Dupoux and colleagues (Dupoux et al. 2001; Rahmani et al. 2015). In that task, speakers whose language includes lexical prosody, like Dutch and Japanese, outperform speakers of languages without, like French.

Languages with postlexical prosody affecting the representation of words, like French (tone), Bengali (stress) and Italian (long vowels), should be distinguished from languages that have no word prosodic representation either lexically or postlexically, like varieties of Malay, including Indonesian (Maskikit-Essed and Gussenhoven 2016). Ambonese Malay has intonational melodies, but they can only be related to the edges of intonational phrases and the final word. Declaratives end in a rising-falling contour, which tends to be completed within the space of the final word, but there is no systematic temporal alignment of the contour with any syllable. The peak will generally be closer to the phrase end if the final word is a monosyllable than if it is polysyllabic. The melody can thus be described as a floating LHL boundary tone complex, but not as a stress-bound or syllable-bound pitch accent.

11.3 DESCRIBING STRESS PATTERNS

Stress is a property of syllables. However, since there was no representation of the syllable at the time, the earliest phonological description indicated stress by employing an n-ary feature [nstress] to be included in the feature specification of vowels. Apart from the questionable attribution of stress to vowels instead of syllables, there were two unsatisfactory aspects of this analysis. First, word stress was conceived of as varying levels of stress across the syllables (vowels) of a word, with [1stress] being the most prominent (Chomsky and Halle 1968). This meant that (a vowel in) a syllable with word stress and a pitch accent was seen as possessing a more intense level of stress than a stressed syllable without a pitch accent, something which most phonologists today will see as a conflation of different concepts, stress and tone. The other unsatisfactory aspect was that stress specifications for syllables ignored the fact that the stress systems of languages reflect the alternation tendency noted at the beginning of this chapter. Stressed syllables typically alternate in a binary fashion, s-w-s-w-s-w, as opposed to, say, s-s-w-s-s-w. However, a rule assigning stress after every two, three, four or more syllables is not more or less complex than a rule assigning stress to every odd syllable. What would be needed here is some representation of stress that imposes, or favours, this ‘counting-by-two’ aspect.

11.3.1 Footing

The first proposal for gathering syllables into groups of two was made by Liberman and Prince (1977) and Vergnaud and Halle (1978), who introduced a binary branching tree with a strong and a weak node, referred to as a ‘foot’ by Selkirk (1978). Adopting this representational device, Hayes (1981) formulated the three
parameters listed in (3) to characterize stress systems. Instead of trees, we will use two-layered foot and word structures to indicate the metrical structure.

(3) Dominance: Iamb vs Trochee  
   Edge: Left to right vs right to left  
   Weight: Quantity-sensitive vs quantity-insensitive

### Dominance

Languages may choose whether the right or left node of a binary tree is dominant. A w-s structure is known as an *iamb* (ˈaɪ.əm(b))/, while a s-w structure is a *trochee* (ˈtrəʊ.ki/ or ˈtrəʊ.kiː/). Many European languages are trochaic, while many American languages are iambic (van der Hulst et al. 2010).

### Edge

Languages differ in whether the main stress occurs near the beginning or the end of the word, expressed by the choice for the second parameter. The words in left-edge languages may, for instance, have either initial or peninitiial stress, while those in right-edge languages may have penultimate or final stress. There are two sub-parameters under 'Edge'. First, when foot structure covers the whole word, the footing is 'iterative', and feet are built until the opposite word edge is reached. This potentially creates two or more stressed syllables in the word. Iterative footing is also known as 'rhythmic stress'. Second, when iterative footing ends up excluding just a single syllable, some languages leave that syllable unfooted (a 'stray' syllable), whereas others turn it into a monosyllabic foot, known as a 'degenerate' foot.

### Weight

In many languages, the composition of feet is sensitive to the segmental composition of syllables. A highly frequent restriction is that a weak node may not dominate a bimoraic, i.e. 'heavy', syllable. If the language builds trochees from the right edge, then a structure like [ta.ta.ta.ta:] would not tolerate a trochee over the last two syllables, since word-final [ta:] would end up in its weak position. Instead, the final syllable would be a monosyllabic foot, with the disyllabic trochee being built over the penultimate and antepenultimate syllables: [ta.’(ta.ta).’(ta:)]. Languages that are quantity-insensitive build the foot structures over their words without regard for the segmental composition of syllables, even in cases in which the language has both heavy and light syllables.

---

Q108  **English builds quantity-sensitive trochees from the right edge of the word. Determine on the basis of the following words whether English has stray word-initial syllables.**
We will illustrate the three parameters with the stress systems of three languages. First, Icelandic words begin with the main stress, and secondary stress occurs on other odd syllables in the word (Árnason 1985). This is shown in (4), where feet are indicated by parentheses, with the stressed syllable marked ‘x’, while ‘.’ indicates an unstressed syllable. To indicate that the first stressed syllable has the main stress, an extra layer is added over the foot which represents the word. In it, the stressed syllable with the main stress has an extra x placed over it.

In terms of these first three parameters, Icelandic builds trochees from the left word edge and is quantity-insensitive. For this last point, notice that closed syllables tolerate being in an unstressed position, as shown by [ðiŋ] in (4c) and [tem] in (4d).

The second and third languages differ from Icelandic and from each other in terms of ‘Dominance’, ‘Edge’, the sub-parameters of ‘Edge’ and ‘Weight’. Garawa shows that the edge which is relevant for the main stress may be different from that for the secondary stresses. This type of language indicates that the main stress may become disassociated from an alternating rhythm, and suggests that it is important to distinguish ‘rhythmic stress’ from ‘main word stress’. Examples (5a,b) show perfect Icelandic structures, but while Garawa builds a trochee at the left word edge for the main stress, secondary stresses occur on even syllables, counting from the right, as shown in (5c), where iteration from the left would have created a trochee [la.rin] (Furby 1974; Hayes 1981). Regardless of the edge from which they are built, the trochees are again quantity-insensitive, as shown by [cim] in (5d) and [rim] in (5e). Examples (5c,e,f) show another difference with Icelandic. Left-over syllables are ‘stray’ in Garawa, but are footed in Icelandic. These degenerate feet were shown in (4c,d,e).
Hopi has a quantity-sensitive iamb at the left word edge. Also, unlike Icelandic and Garawa, it doesn’t appear to have rhythmic stress, and thus builds its foot non-iteratively. All syllables outside the left-edge foot are thus ‘stray’.

Q109

1 Why can’t Hopi be described as having a left-edge or right-edge trochee?
2 Hopi [ca.qap.ta] ‘disk’ has the main stress on the second syllable. This shows it is quantity-sensitive. True or false?

A fourth parameter proposed by Hayes (1981), ‘Boundedness’, allows the characterization of stress systems which fail to show the typically observed tendency to confine the location of the main stress to a small domain near a word edge. Such unbounded languages don’t have rhythmic stress and are always quantity-sensitive. For instance, main stress in monomorphemic words in Amele is on the leftmost
heavy (i.e. closed) syllable, and if there are no heavy syllables, it is on the initial syllable. A diphthong in word-final position counts as heavy, but as light elsewhere (Roberts 1987: 357). The condition without heavy syllables is illustrated in (7a,b,c). Examples (7d,e,f,g) each contain one heavy syllable, which attracts the stress. Roberts (1987) doesn’t give any nouns or adjectives with two heavy syllables; (7h) is a verb illustrating leftmost choice out of two heavies. No foot structure is indicated, since the pattern was derived without making a choice between trochee and iamb.

(7) a ‘ma.la ‘chicken’
b ‘meu.la ‘right hand’
c ‘ni.fi.la ‘k.o. beetle’
d bai.’al ‘wooden sword’
e ‘ʔo.’o.’wai ‘praying mantis’
f ja.’wai.ti ‘north wind’
g ‘laʔ.ɡbai.a ‘scorpion’
h ‘wol.doʔ ‘surpass’

Amele has leftmost default stress and stresses the leftmost heavy syllable, but the specification of the edge of the stress when no heavy syllables are present and the rightmost/leftmost specification of the heavy syllables are independent. Thus, four types of unbounded systems have been reported (Baković 2004). Gordon (2000) observes that there is a paucity of phonetic work, in particular on languages with a ‘default-to-opposite-edge’ pattern, and that more research is called for.

11.3.2 Quantity-sensitivity biases for trochees and iambs

A parametric theory carries an expectation of orthogonal dimensions, such that any combination of selected options is readily attested. Unbounded languages go against this expectation, because of the implication that in such languages quantity-sensitivity and absence of rhythmic stress go hand in hand. Hayes (1985, 1987) further pointed out that iambs are overwhelmingly quantity-sensitive. He connected this with the results of perception experiments which show that listeners group sequences of stimuli of equal duration in a binary fashion with initial prominence, i.e. (tata) (tata) (tata), but sequences of stimuli with unequal duration in binary groups with final prominence, i.e. (tataa) (tataa) (tataa), the Iambic/Trochaic Law. This suggests that iambic rhythm is expressed durationally. To capture the typological bias, he introduced the foot inventory in (8), where σ is either μμ or μ. The syllabic trochee is meant to cater to quantity-insensitive systems, and cannot be instantiated by a monosyllable. The inherently quantity-sensitive iamb and the moraic trochee should cover quantity-sensitive systems, and both of these can be monosyllabic. For the disyllabic version of the moraic trochee we deviate from Hayes (1995) in allowing a heavy strong branch, (μμ μ), by the side of (μ μ), instead of just the ‘even’ moraic trochee [μ μ].

(8) a SYLLABIC TROCHEE σ σ
    b MORAIC TROCHEE σ μ if possible, or else μμ
    c IAMB μ σ if possible, or else μμ
Q110 Cairene Arabic has iterative footing from the left, with main stress on the final foot.

1 Why do the positions of the main stress in items (a) and (b) differ from those in items (c), (d) and (e)?

2 Do these data support Hayes's position to exclude the 'uneven' trochee [μμ μ]?

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>(ˈka.ta).ba</td>
<td>'he wrote'</td>
</tr>
<tr>
<td>b</td>
<td>(ˌfa.ga).('ra.tu).hu</td>
<td>'his tree NOM'</td>
</tr>
<tr>
<td>c</td>
<td>(ˌfa.ga).(ˌra.tu).(ˈhu.ma)</td>
<td>'their dual tree NOM'</td>
</tr>
<tr>
<td>d</td>
<td>(ˌʔad).(ˌwi.ya).(ˈtu.hu)</td>
<td>'his drugs NOM'</td>
</tr>
<tr>
<td>e</td>
<td>(ˌmər).(ˈta.ba)</td>
<td>'mattress'</td>
</tr>
</tbody>
</table>

11.3.3 Flexible right edges

Stress systems may behave as if the right edge of the word in fact occurs before the physical word end. It is as if some final element, the last consonant, mora or syllable, is systematically ignored, in which case the ignored element is said to be extrametrical. A subpart of the Turkish lexicon, among which are underived place names, behaves as if the final syllable is extrametrical. The position of the main stress is on the antepenult if it is heavy (equivalent to ‘closed’, since Turkish lacks vowel quantity) and the penult is light, but otherwise is on the penult (Sezer 1983). The fact that only the penult and the antepenult are involved in the generalization suggests we should ignore the final syllable and think of how the two syllables that do matter could be grouped into a binary foot. Examples (9a,b) illustrate how stress goes to the closed antepenult if the penult is open, and (9c,d) show how an open antepenult forces the stress to occur on the penult. The antepenult in (9e,f) is closed, but since the penult is also closed, stress is on the penult. In all cases, the final syllable varies between closed, as in (9a,c,e), and open, as in (b,d,f), without effect.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>ˈer.zu.rum</td>
<td>‘Erzurum’</td>
</tr>
<tr>
<td>b</td>
<td>ˈan.ka.ra</td>
<td>‘Ankara’</td>
</tr>
<tr>
<td>c</td>
<td>in.ˈdi.a.na.ˈpolis</td>
<td>‘Indianapolis’</td>
</tr>
<tr>
<td>d</td>
<td>a.ˈla.ˈba.ma</td>
<td>‘Alabama’</td>
</tr>
<tr>
<td>e</td>
<td>is.ˈtæn.bul</td>
<td>‘Istanbul’</td>
</tr>
<tr>
<td>f</td>
<td>an.ˈtal.ja</td>
<td>‘Antalya’</td>
</tr>
</tbody>
</table>

Another theoretical device that has been proposed is catalexis (Kiparsky 1991). It specifies an empty final syllable (or mora) which is seen by the footing algorithm but is not otherwise observable.

Q111 Assuming that the final syllables of the words in (9) are extrametrical, do these data require a moraic trochee or an iamb to account for the position of the main stress?
11.3.4 Morphological effects and exceptions

Morphological processes frequently impose specific stress locations, which is why accounts are usually based on the evidence of underived words. In English, for instance, there is a group of suffixes which have this property, like -ee, -eer, -ic, -ity, as illustrated by nomi'nee (cf. 'pedigree'), mountai'neer (cf. 'Windermere), or'ganic (cf. 'limerick') and fri'gidity (cf. Missi'sippi). There are, broadly, two ways in which such effects have been accounted for. One is to use the regular stress assignment algorithm of underived words, but provide 'stress-changing' suffixes with a feature like [±cyclic], as in Halle and Kenstowicz (1991). The other way is to specify the foot structure they impose on the base in their underlying representation, along with the segments, as standardly assumed for Slavic (Stankiewicz 1993) and applied to English in Gussenhoven (1994). A ‘pre-accenting’ suffix like -ity would come with a trochee with its weak position over the penultimate syllable and its strong position over an empty syllable before it, while -ese comes as a monosyllabic trochee. Exceptions can similarly be treated in two ways. One option is for the lexical representations to include a property that is recognized by the theoretical framework. Italian, for instance, has regular penultimate stress but has a number of underived words with antepenultimate stress as well as words with final stress. Its penultimate regular stress can be seen in loanword adaptation, as in the case of German ['ham bourk], which has been taken over as [am 'burgo] (Sluyters 1990), where penultimate stress is obligatory because of the closed syllable. The system is thus readily described with the help of a moraic trochee aligned on the right. To account for the exceptional words like [ˈka.li.tʃe] 'chalice', a final extrametrical syllable can be listed in the lexicon (cf. regular [kɔr ni.tʃe] 'frame'), while words like [tʃit. ta] 'city' could have a final catalectic syllable. The alternative is to provide their lexical entries with the foot structure concerned. The rule Sezer (1983) gave for the word group exemplified in (9) in fact has many exceptions, in addition to variation (e.g. [ˈyːs.ky.dar] or [ys.'ky.dar] 'Üsküdar').

Q112 A possible proposal for Turkish is to simply prespecify stressed syllables or feet in the lexicon (cf. Barak and Vogel 2001; Özçelik 2014).

1 Prespecifying the metrical structure of all ‘Sezer’ words goes against a widely observed phenomenon in stress systems. What is it?
2 Which of the following data are exceptions to the description in Sezer (1983)?
   a [bel.'tʃi.kə] 'Belgium'
   b [ams.'ter.dam] 'Amsterdam'
   c [ys.'ky.dar] 'Üsküdar'
   d [er.'zin.can] 'Erzincan'
   e ['ko.pe.nag] 'Copenhagen'
3 A possible objection against a specification of metrical structure in the lexicon for exceptional words or affixes is that structures can be specified that do not exist in the language, like a suffix imposing word-initial main
stress in English words. By accounting for exceptionality and morphological effects in terms of theoretical devices that are more restrictive, like catalexis or extrametricality, such overgeneration could be avoided. Can you think of an argument to neutralize this objection?

11.3.5 Independent evidence for the foot

Ideally, the reality of the feet that are constructed to account for the stress locations in words is also revealed by the way in which other generalizations refer to them. Such confirmation is in principle provided by all languages in which different segmental conditions hold for stressed and unstressed syllables, like English (Harris 2013). In addition, there may be phonological and phonetic processes which refer to the foot. English vocative chant, which consists of a higher and a lower pitch, analysed as the H*IH pitch accent, spreads its first tone until it reaches a stressed syllable or the final syllable of the intonational phrase. In (10a), which has a secondary stress on the penult, the first pitch level covers *cater-*, while the second covers *pillar*. In (10b), the main stress is again on the pre-antepenult, but because the secondary stress is word-final, the first level covers the three syllables *catama-*. Example (10c) shows that if there are no secondary stresses, the second level also begins on the last syllable.

!(10)
\[
\begin{align*}
a & \quad [ˈkæ.təˌpi.lə] \quad \text{‘caterpillar’} \\
b & \quad [ˈkæ.tə.məˌræn] \quad \text{‘catamaran’} \\
c & \quad [ˈkɔː.mə.ɾənt] \quad \text{‘cormorant’}
\end{align*}
\]

Q113 English [ŋ] obligatorily appears before dorsal consonants in the words in (1), but varies with [n] in the words in (2) (Kiparsky 1979). What is the domain of the obligatory occurrence of [ŋ]? Motivate your answer.

(1) ˈɪŋk \quad \text{ink}  \\
     ˈɪŋ.kra.mənt \quad \text{increment}  \\
(2) ɪŋ.ˈkriːs \quad \text{or} \quad ɪŋ.ˈkriːs \quad \text{increase} (\text{verb})  \\
     ɪŋ.ˌkriːs \quad \text{or} \quad ɪŋ.ˌkriːs \quad \text{increase} (\text{noun})

Q114

1 Dutch has a vocative chant which is similar to the English tune. In [ˈnɔːt. por.ˌtiːr] ‘night porter’, a compound of [ˈnɔːt] and [por.ˌtiːr], the first pitch level includes the second syllable, while in [ˈnɔːt.ˈvaː.ˌkɔːr] ‘night
Watchman, a compound with a second constituent [‘va:.kor], it does not. What metrical representation does this suggest for the first syllable of [por. ‘ti:ri]?  

2 Dutch [ﬁ] is pronounced in the words in (1), but, despite the spelling, there is no [ﬁ] in the words in (2). What determines its distribution?

(1) ['fiut] hoed 'hat'  
['a:.bra.ˌnæm] Abraham 'Abraham'  
[jo.ˈtænas] Johannes 'John'  
[fio.ˈtel] hotel 'hotel'

(2) ['ni.o] Niehe (proper name)  
[a.ˈlo:a] aloha 'aloha'  
[ˈta:.u] tahu 'tofu'

**Q115** Yabem, a tone language, has stress on the last syllable, as shown by the data below (Ross 1993).

<table>
<thead>
<tr>
<th>a</th>
<th>wá</th>
<th>'mango'</th>
<th>k</th>
<th>pá.ˈliŋ</th>
<th>'careless'</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>wà</td>
<td>'crocodile'</td>
<td>l</td>
<td>bá.ˈliŋ</td>
<td>'far away'</td>
</tr>
<tr>
<td>c</td>
<td>ó.ˈlí</td>
<td>'body'</td>
<td>m</td>
<td>sá.ˈkíŋ</td>
<td>'service'</td>
</tr>
<tr>
<td>d</td>
<td>ó.ˈlì</td>
<td>'wages'</td>
<td>n</td>
<td>sá.ˈɡíŋ</td>
<td>'house partition'</td>
</tr>
<tr>
<td>e</td>
<td>á.ˈwè</td>
<td>'outside'</td>
<td>o</td>
<td>ká.ˈká.ˈtòŋ</td>
<td>'I made a heap'</td>
</tr>
<tr>
<td>f</td>
<td>à.ˈwè</td>
<td>'woman'</td>
<td>p</td>
<td>ká.ˈlè.ˈtí</td>
<td>'I ran'</td>
</tr>
<tr>
<td>g</td>
<td>jà.ˈó</td>
<td>'prohibition'</td>
<td>q</td>
<td>ká.ˈkí.ˈlì</td>
<td>'I provoked trouble'</td>
</tr>
<tr>
<td>h</td>
<td>jà.ˈô</td>
<td>'enmity'</td>
<td>r</td>
<td>ká.ˈlè.ˈsù</td>
<td>'I poked'</td>
</tr>
<tr>
<td>i</td>
<td>kúŋ</td>
<td>'call out'</td>
<td>s</td>
<td>ká.ˈɡé.ˈlì</td>
<td>'I stepped over s.t.'</td>
</tr>
<tr>
<td>j</td>
<td>gün</td>
<td>'spear s.t.'</td>
<td>t</td>
<td>ká.ˈmá.ˈdòm</td>
<td>'I broke in two'</td>
</tr>
</tbody>
</table>

1 Determine the Yabem foot in terms of foot type and word edge.  
2 Tones do not freely vary from syllable to syllable. What constrains their freedom?  
3 Is voicing in plosives contrastive?  
4 What is the 1sg morpheme?  
5 What are the Yabem translations of 'I called out' and 'I speared something'?

**11.4 STRESS AND OPTIMALITY THEORY**

More recent descriptions of stress systems are frequently presented within Optimality Theory. Recall that this theoretical programme is committed to isolating structural properties as single constraints. In the case of stress, this commitment may not
seem to sit well with the observation that structural properties frequently correlate, as illustrated by the Iambic/Trochaic Law, for instance. Hayes’s Iamb in (8) combines a dominance property and the property of being quantity-sensitive, but a fine-grained approach would encode dominance and quantity-sensitivity in separate constraints. In this section, we will lean towards fine-grained versions of the various constraints, as this will show best how this works. Another feature of Optimality Theory which will be relevant to stress is its unselective definition of the input. It should readily be capable of dealing with any phonological structure, however overspecified or underspecified. This feature is known as Richness-of-the-Base, often abbreviated as ROTB, and implies that any input will end up as some well-formed output. This makes the lexical specification of metrical structure for words with exceptional stress unproblematic in principle, and thus renders the inclusion of diacritic features like [+cyclic] less desirable. A third aspect of Optimality Theory which will be brought out in the discussion is its working hypothesis that constraints are grounded in phonetics, a goal that was earlier expressed with respect to distinctive features in Postal’s (1968) Naturalness Condition and, perhaps less successfully, with respect to phonological rules (‘processes’) in Stampe (1980).

11.4.1 Foot shapes: dominance, quantity-sensitivity and degenerate feet

Dominance is taken care of by ranking the constraints (11a) and (11b). As will be clear, right-dominance results if (11a) outranks (11b), left-dominance if the ranking is the reverse.

(11)  
   a  IAMB: Feet are iambic.
   b  TROCHEE: Feet are trochaic.

The foot constraints in (11) could be made more elaborate, so as to reflect the statistical tendencies expressed in (8), as in the Uneven Iamb (Kager 1999: 151), which is satisfied by (ta.’ taa), but not by (ta.’ ta) or (’ta), which fail to have a moraic strong position, and not by (’taa), which lacks the initial light position. For (11), we assume the fine-grained interpretation that they just ban feet showing the reverse dominance. That is, (11a) is violated by (’taa), (’taa), (’taa) and (’taa), but satisfied by (’ta), (’taa), (ta.’ taa), (ta.’ ta) and (tad. taa) as well as by the absence of a foot. Whereas the direction of dominance is taken care of by the relative ranking of the constraints in (11), foot construction and binary rhythm are taken care of by constraints (12) and (13).

(12)  Parse-σ: A syllable must be parsed into a foot.
(13)  Foot Binary (FrBin): A foot must be binary (either two syllables or two moras).

These two constraints will lead to an alternation between weaker and stronger articulations of syllables, as illustrated in (14) for marmolada ‘marmelade NOM; sg’ in the quantity-insensitive trochaic language Polish (Dogil and Williams 1999). Candidates (c) and (d) are ruled out because not all syllables have been parsed into a foot. Candidates (b) and (e) are ruled out because they contain feet that are not binary.
Quantity-sensitive languages require the presence of stress either on long vowels or on both long vowels and closed syllables. This property is enforced by (15), which disallows a bimoraic syllable in the weak position of a foot.

(15) **Weight-to-stress Principle (WSP):** Bimoraic syllables are stressed.

The inclusion of closed syllables among the bimoraic syllables of a language is effected by constraint (16). The term ‘Weight-by-Position’ refers to a consonant’s entitlement to moraic status on the grounds of its position in the coda as opposed to the onset (Hayes 1989a).

(16) **Weight-by-Position (Coda-μ):** A coda consonant is moraic.

The ranking of Parse-σ relative to WSP and Coda-μ determines whether a language is quantity-sensitive. Closed syllables in Polish do not attract stress, as shown by the fact that disyllabic words with final open syllables, like *cawa* [ˈka.va] ‘coffee’, have the same position of main stress as disyllabic words with final closed syllables, like *Przemek* [ˈpʃɛ.mɛk] ‘boy’s name’. In Spanish, prosodically similar words, like *casa* [ˈka.sa] ‘house’ and *papel* [paˈpel] ‘paper’, have different stress locations, due to the language’s sensitivity to the bimoricity of closed syllables. The difference is captured by ranking Parse-σ below WSP and Coda-μ for Spanish, as in (17), but above these two constraints for Polish, as in (18).

Thus, to characterize a quantity-insensitive language like Polish or Icelandic, Parse-σ must rank above WSP and Coda-μ. The ranking of the latter two may seem...
immaterial, as can be established by reversing the ranking of WSP and CODA-µ in (18). However, with a higher-ranking Parse-σ, the ranking WSP » CODA-µ characterizes closed syllables as monomoraic rather than bimoraic, and if we wish to suppress unmotivated bimoricity for closed syllables, as indeed we should, WSP must outrank CODA-µ. This same ranking equally characterizes iambic quantity-insensitive languages, like Shiwi (Valenzuela and Gussenhoven 2013). Again, Parse-σ » WSP ensures that a disyllabic foot is formed, on the left edge as it happens, while WSP » CODA-µ ensures that closed syllables are monomoraic, as shown in (19) for the word for ‘manioc stick’, where high-ranking Iamb is assumed.

(19)  

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kən.</td>
<td>ɲa.</td>
<td>la</td>
</tr>
<tr>
<td>a</td>
<td>µ µ µ (kən. ɲa).la</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>b</td>
<td>µ µ µ (ˈkən).ɲa.</td>
<td>la</td>
<td>**!</td>
</tr>
<tr>
<td>c</td>
<td>µ µ µ (kən.ˈɲa).la</td>
<td>*</td>
<td>*!</td>
</tr>
</tbody>
</table>

The constraint Parse-σ interacts not only with WSP (15) and CODA-µ (16) to express whether a language is quantity-sensitive or quantity-insensitive, but also with FtBin (13) so as to account for a language’s rejection (Spanish, Shiwi) or acceptance (Icelandic) of degenerate feet. To prevent the final syllable of [kən.ˈɲa.|la] from forming a monomoraic foot, FtBin must outrank Parse-σ. The effect is shown in (20).

(20)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>kən.</td>
<td>ɲa.</td>
</tr>
<tr>
<td>a</td>
<td>µ µ µ (kən. ɲa).la</td>
<td>*</td>
</tr>
<tr>
<td>b</td>
<td>µ µ µ (kən.ˈɲa).ˈla</td>
<td>*!</td>
</tr>
</tbody>
</table>

The reverse ranking is required for rightward trochaic Icelandic as well as for leftward iambic Weri, both of which will parse left-over syllables as degenerate feet (Boxwell and Boxwell 1966). This is illustrated for the Weri word ‘hair of arm’ in (21). In Icelandic, the same ranking will reject [(ˈseptem)ber] in favour of [(ˈseptem)(ber)], assuming high-ranking Trochee.

(21)  

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ku.</td>
<td>li.</td>
</tr>
<tr>
<td>a</td>
<td>µ µ µ (ˈku).li.</td>
<td>pu</td>
</tr>
<tr>
<td>b</td>
<td>µ µ µ ku. (li.</td>
<td>pu)</td>
</tr>
</tbody>
</table>

Quantity-sensitive languages must rank Parse-σ below WSP and CODA-µ. Observe, contrary to (20a) for Shiwi, how an initial disyllabic iamb for the Hopi word for ‘roofbeam’ (6a) is rejected because it fails to obey WSP for the initial
stressed syllable, as shown in (22). The same result is obtained for languages without closed syllables but with a quantity contrast for vowels, like Hawaiian, which is quantity-sensitive for long vowels and diphthongs, as in (23) for the geographical name Ni‘ihau.

(22) | les.ta.vi | WSP | CODA-µ | PARSE-σ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>µ µ µ (les. ‘ta).vi</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b</td>
<td>µµ µ µ (les. ‘ta).vi</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>c</td>
<td>µµ µ µ (les).ta_.vi</td>
<td></td>
<td>**</td>
</tr>
</tbody>
</table>

(23) | ni.ʔi.hau | WSP | CODA-µ | PARSE-σ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>µ µ µ µ ni.(ʔi.hau)</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>b</td>
<td>µ µ µ µ ni.ʔi.(ʔau)</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>c</td>
<td>µ µ µ µ (ni.ʔi).hau</td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>

The ranking of WSP above PARSE-σ thus describes the quantity-sensitivity of long vowels and diphthongs. The relative ranking of CODA-µ and PARSE-σ is responsible for treating closed syllables as heavy or as light. In (22), the bimoraic closed syllable results from ranking CODA-µ above PARSE-σ. Let us next consider how the fine-grained approach encoding dominance and quantity-sensitivity in separate constraints allows us to describe quantity-sensitive languages which treat closed syllables as light, like Lenakel. It has a right-edge quantity-sensitive trochee and rhythmic stress from the right, as illustrated in (24), from Hayes (1995). As illustrated in (25), this requires that WSP dominates PARSE-σ, necessary for stressing a final syllable containing a long vowel, but where PARSE-σ itself dominates CODA-µ, in order to prevent stressing a final closed syllable, as illustrated in (26).

(24) | ‘na.puk | ‘song’ |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>‘ti.ˈkom.kom</td>
</tr>
<tr>
<td>b</td>
<td>tu.p* a.lu.’ka.luk</td>
</tr>
<tr>
<td>c</td>
<td>ˌri.me.ˈtjaːw</td>
</tr>
<tr>
<td>d</td>
<td>‘ti.ˈna.ˌkam.wap.ˈkeːn</td>
</tr>
</tbody>
</table>

(25) | ri.me.tjaːw | WSP | PARSE-σ | CODA-µ |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>µ µ µ µ (‘ri.me).ˈtjaːw)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>µ µ µ µ (‘ri.me).tjaːw)</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>c</td>
<td>µ µ µ µ (‘ri.me).tjaːw)</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>d</td>
<td>µ µ µ µ (‘ri.me).tjaːw)</td>
<td>*!</td>
<td>*</td>
</tr>
</tbody>
</table>
Q116
1 There are no quantity-sensitive languages in which long vowels are light and closed syllables heavy. How does that follow from the analysis?
2 Bengali has quantity-insensitive word-initial stress, with binary rhythmic stress on every odd syllable in the word, rejecting degenerate feet (see section 11.2.3). It also has geminates, as in [ˈta.ləb.bo] 'palatal'. Assuming high-ranking TROCHEE and a moraic representation of geminates, can this language be described by a ranking of WSP, FTBIN and PARSE-σ?

11.4.2 Edges and iteration
McCarthy and Prince (1993, 1998) capture the left-to-right or right-to-left parsing of feet by appealing to alignment, a device to locate phonological or morphological constituents in the phonological structure. Recall from chapter 4 how the position of an affix is determined by a constraint which aligns one of its edges with the correct edge of the word, left for prefixes and right for suffixes. Generally, an alignment constraint stipulates that the right or left edge of a constituent (root, stem, morphological word, mora, syllable, foot, phonological word (ω), intonational phrase, etc.) must coincide with the right or left edge of some other constituent. Alignment constraint (27a) requires that the right edges of any phonological word and a foot should coincide, while (27b) does the same for the beginning of the phonological word. If (27a) outranks (27b), a trisyllabic word will be parsed as σ(σσ), while the opposite ranking leads to (σσ)σ. As will be clear, instead of ω, languages may also position feet relative to morphological constituents, like the morphological word.

(27) a ALIGN(ω, Foot, R): The right edge of a phonological word is aligned with the right edge of a foot.
b ALIGN(ω, Foot, L): The left edge of a phonological word is aligned with the left edge of a foot.
Hopi, which as we have seen has an iamb at the left word edge, is described by ranking the two constraint pairs (11) and (27) as \textsc{iamb} $\gg \textsc{trochee}$ and \textsc{Align}(\omega,\textsc{Foot},\textsc{L}) $\gg \textsc{Align}(\omega,\textsc{Foot},\textsc{R})$. Since the iamb is quantity-sensitive, both WSP and \textsc{coda-\textmu} must dominate \textsc{parse-\sigma}. Violations of (27) are indicated by the number of syllables each foot is removed from the relevant word edge, the latter specified by \#, as shown in (28).

\begin{verbatim}
(28) \hline
  & les.ta.vi & \textsc{iamb} : Al(\omega,Ft,L) & WSP & \textsc{coda-\textmu} & \textsc{parse-\sigma} & \textsc{trochee} : Al(\omega,Ft,R) \\
\hline
  a & µµµµ(‘les.ta.vi) & & & & & \#σσσ\# \\
  b & µµµµ(‘les.ta.vi) & *! & & & * & α# \\
  c & µµµµ(les.(‘ta.vi) & #σ*! & *! & & * & * \\
  d & µµµµ(les.‘ta.vi) & *! & * & * & * & α# \\
  e & µµµµ(les.‘ta.vi) & *! & * & * & * & α# \\
\hline
\end{verbatim}

Hopi doesn’t have iterative stress, but many languages of course do. Iterative footing can be enforced by demanding that feet align with a word edge, as opposed to demanding that a word edge align with a foot edge. This constraint pair is given in (29).

\begin{verbatim}
(29) \hline
  a & \textsc{align}(\textsc{foot},\omega,\textsc{r}) & The right edge of every foot is aligned with the right edge of the phonological word. \\
  b & \textsc{align}(\textsc{foot},\omega,\textsc{l}) & The left edge of every foot is aligned with the left edge of the phonological word. \\
\hline
\end{verbatim}

The number of violations of (29) is equal to the sum of the syllables that each foot is removed from the relevant word edge. To minimize those violations, \textsc{parse-\sigma} (12), which generally promotes exhaustive footing, will have to be ranked below (29), as it must be in non-iterative Hopi. The reverse ranking give iterative footing, as tableau (30) illustrates for the Garawa word for ‘wrist’, after Kenstowicz (1994b). Recall from section 11.3.1 that Garawa is quantity-insensitive, is trochaic, has rhythmic stress from the right and disallows degenerate feet. \textsc{parse-\sigma} must be ranked below FrBin to prevent degenerate feet, but above (29) to enforce rhythmic stress. Thus, candidates (30a,b) each incur a single violation of \textsc{parse-\sigma}, but only (30a) correctly aligns its left word boundary with a foot. Candidate (30c) fails to have any rhythmic footing, while (30d) is too generous in its parsing, incurring a violation of FrBin.

\begin{verbatim}
(30) \hline
  & ka.ma.la.rin.ji & FrBin & \textsc{parse-\sigma} & Al(\omega,Ft,L) & Al(\omega,Ft,R) & Al(\omega,Ft,L) \\
\hline
  a & (‘ka.ma).la.(‘rin.ji) & * & & σσσ# & & #σσσ \\
  b & ka.(‘ma.la).(‘rin.ji) & * & #σ! & σσ# & & #σ, #σσσ \\
  c & (‘ka.ma).la.rin.ji & **!* & σσσ# & σσσ# & & #σ, #σσσ \\
  d & (‘ka).l.(‘ma.la).(‘rin.ji) & *! & & σσσ# & σσσ# & #σ, #σσσ \\
\hline
\end{verbatim}
Q117

1 All syllables of the Garawa word [nim.paləŋi.mun.kun.了一句 mi.ɾa] ‘from our own two’ are parsed into feet. Why?
2 Using a tableau with the constraints as ranked in (30), show the fates of the following output forms for the Garawa word for ‘fought with boomerangs’:
   a (ˈŋan.ki)(ˌri.ka)(ˌrim.pə).ji
   b (ˈŋan.ki)(ˌri.ka).rim.(ˌpa.ɾi)
   c (ˈŋan.ki).ri.ɺ(ˌki.ɾim).ɺ(ˌpa.ɾi)

In (30), rhythmic stress, due to $A_L(Ft,ω,R)$, represents the secondary stress, with the main stress being due to $A_L(ω,Ft,L)$. The implication that rhythmic stress is secondary stress when the main stress and rhythmic stress refer to different edges is not enshrined in Optimality Theory. A third pair of alignment constraints specifies the edge where the main stress is found. In (30), the understanding is that (31b) outranks (31a).

(31) a $\text{ALIGN(Head,}ω,R)$: The foot with main stress is aligned with the right edge of the phonological word.
   b $\text{ALIGN(Head,}ω,L)$: The foot with main stress is aligned with the left edge of the phonological word.

Q118 Cavineña (Key 1968) has quantity-insensitive stress with the main stress on the right.

1 Is the language trochaic or iambic?
2 Draw a tableau with the four constraints $\text{PARSE-σ}$, $\text{FtBIN}$, $\text{AL(Ft,}ω,L)$ and $\text{AL(Ft,}ω,R)$. Your tableau should have dotted lines between columns representing unranked constraints. Illustrate your analysis with the word [a.ta.ta.wa.ha].

   a ‘e.na ‘water’
   b ki.ˈɾi.ka ‘paper, book’
   c ,a.si.ˈka.da ‘dirty’
   d a.ˌta.ta.ˈwa.ha ‘kind of bee’
   e ,a.ta.ˌta.wa.ˈha.ke ‘kind of bee+poss’
Q119 Creek has the main stress on the right, preceded by rhythmic stress (Haas 1977; Hayes 1995).

1 Identify the forms which reveal whether Creek is iambic or trochaic.

2 Draw a tableau with the five constraints Parse-σ, Al\(\text{Ft}, \omega, R\), Al\(\text{Ft}, \omega, L\), WSP and Coda-\(\mu\), using item (f) with candidates \([, \text{al}^\mu\].\text{pa.} \text{'to}.\text{ci}\) (where superscript \(\mu\) indicates a moraic coda), \([, \text{al}].\text{pa.} \text{'to}.\text{ci}\), \([, \text{al}^\mu\].\text{pa.} \text{'to}.\text{ci}\), \([, \text{al}^\mu\].\text{pa.} \text{'to}.\text{ci}\) and \([, \text{al}^\mu\].\text{pa.} \text{'to}.\text{ci}\) as an illustration.

|a| i.‘fa | ‘dog’ |
|b| i.‘fo.ci | ‘puppy’ |
|c| po.‘cos.wa | ‘axe’ |
|d| a.‘pa.ta.‘ka | ‘pancake’ |
|e| ‘ak.to.‘pa | ‘bridge’ |
|f| ‘al.pa.‘to.ci | ‘baby alligator’ |
|g| ‘naf.ki, ti.ka., i.‘ji.‘ta | ‘hit-PL;OBJ’ |

11.4.3 Geminates and weight

The moraic representation of geminates (see chapter 9 and Q116(2)) predicts that quantity-sensitive languages with geminates in postvocalic position, as in Italian [‘fat.to] ‘done’, have heavy closed syllables, in addition to heavy long vowels if they have a vowel quantity contrast. Somewhat surprisingly, this does not appear to be true for a great many such languages (Davis 2011). There are a number of responses to this state of affairs. We could liberalize the representation of geminates, such that those making heavy syllables have a moraic status, while those that don’t have a direct double association with two syllable nodes, such that geminates bypass the mora level in both syllables. Another option is to make the weight of a mora dependent on its segmental features. This solution would be parallel to that for TBUs, which in languages like Japanese are tone-bearing only if the segment they are associated with is \([+\text{son}]\), thus excluding obstruents.

Q120 Wolof has initial stress, but peninitial stress if the peninitial syllable has a long vowel and the initial syllable does not, as shown in items (a) to (i) (Bell 2003).

|a| ‘wo.ne.wu | ‘to show off’ |
|b| ‘baa.si | ‘couscous’ |
|c| ‘fee.sal | ‘to blacken’ |
|d| ‘pe.tax | ‘pigeon’ |
|e| ‘jar.goŋ | ‘spider’ |
11.4.4 Non-finality and clashes

In section 10.7.2, we saw that tonal targets tend to be retracted in the proximity of upcoming pitch accents or phrase boundaries. Stress systems reveal phonological reflexes of this behaviour in a tendency to avoid word-final stress as well as adjacent stresses. The corresponding constraints are NonFinH and NoClash.

The insertion of a pitch accent is more problematic in a word-final syllable than in other syllables. Phrase-final pronunciations, which include citation pronunciations, will have only one syllable available for the pitch accent plus any final boundary tones. There are two tendencies in languages which offer relief to such tonal crowding on final syllables. One is the tendency for word-final syllables to be longer (Zhang 2002). English, for instance, liberally allows trimoraic syllables in word-final position (e.g. [ˈkraʊn] crown), but much less so in word-internal position (as in [ˈskoʊndrəl] scoundrel), and such restrictions may be absolute in other languages. The second is for languages to avoid main stress on the word-final syllable. The earlier introduction of extrametricality as a descriptive device thus receives a straightforward motivation. Iambic languages frequently have initial stress in disyllables, where final stress would be expected, as in Hopi and Shiwi lu, as shown for the latter in (32).

\begin{tabular}{ll}
(a) & mi.ˈka.ra.waʔ ‘turkey’ \\
(b) & ˈsi.ssk ‘porcupine’ \\
(c) & i. kər ‘hurt’ \\
(d) & ˈmu.tuʔ ‘head’ \\
(e) & ˈlək ‘1SG’ \\
(f) & i.ˈkər.mu.tuʔ.ˈlək ‘I have a headache’
\end{tabular}

The constraint which can prevent final stress is (33).

(33) \textbf{Non-Finality Head (NonFinH): A foot head may not be final.}

As a quantity-insensitive left-edge iambic language, Shiwi lu has high-ranking $A_L(\omega,Ft,L)$ and $\textsc{iamb}$, so as to create regular peninitial stress (34), but by ranking NonFinH above $A_L(\omega,Ft,R)$ and Parse-$\sigma$, initial stress results in disyllables, as in (35), just as it does in monosyllables, thanks to $A_L(\omega,Ft,L)$, as in (36). Because there

---

1. Is Wolof quantity-sensitive? Motivate your answer.
2. Is Wolof trochaic or iambic?
3. The word for ‘understanding’ is [deg.ˈgʊʊ^.] Is this unexpected?

---
is no rhythmic stress, $\text{ALIGN(Ft,}\omega,R)$ outranks $\text{PARSE-}\sigma$. Form (34c), incidentally, would also be eliminated by $\text{FtBin}$, provided it ranks below $\text{NonFinH}$.

$$
\begin{array}{|c|c|c|c|c|}
\hline
\text{Form} & \text{Align(Ft,}\omega,L) & \text{IAMB} & \text{NonFinH} & \text{Align(Ft,}\omega,R) & \text{Parse-}\sigma \\
\hline
\text{a} & (\text{i.ker.mu.tu?} \text{lak}) & \text{si} & \text{si} & \text{si} & \text{***} \\
\text{b} & (\text{i.ker}.(\text{mu.} \text{tu?)} \text{lak}) & \text{si} & \text{si} & \text{si} & \text{*} \\
\text{c} & (\text{i.ker.mu.tu?} \text{lak}) & \text{si} & \text{si} & \text{si} & \text{****} \\
\hline
\end{array}
$$

$$
\begin{array}{|c|c|c|c|c|}
\hline
\text{Form} & \text{Align(Ft,}\omega,L) & \text{IAMB} & \text{NonFinH} & \text{Align(Ft,}\omega,R) & \text{Parse-}\sigma \\
\hline
\text{a} & (\text{si} \text{ssek}) & \text{si} & \text{si} & \text{si} & \text{*} \\
\text{b} & (\text{si} \text{.ssek}) & \text{si} & \text{si} & \text{si} & \text{!} \\
\text{c} & (\text{si} \text{. ssek}) & \text{si} & \text{si} & \text{si} & \text{!} \\
\hline
\end{array}
$$

Although $\text{NonFinH}$, formulated as in (33), directly reflects the phonetic motivation behind it, a constraint banning word-final feet altogether is required for a language like Macedonian, where the main stress invariably is on the antepenultimate syllable (‘type 4 dialects’; cf. Baerman 1999), as illustrated in (38). By ranking $\text{NonFinF}$ high, the final two syllables are unavailable for stress.

$$
\begin{array}{|c|c|c|c|c|}
\hline
\text{Form} & \text{Align(Ft,}\omega,L) & \text{IAMB} & \text{NonFinH} & \text{Align(Ft,}\omega,R) & \text{Parse-}\sigma \\
\hline
\text{a} & (\text{\textquoteleft \textquoteleft lak}) & \text{si} & \text{si} & \text{si} & \text{!} \\
\text{b} & (\text{\textquoteleft lak}) & \text{si} & \text{si} & \text{si} & \text{!} \\
\text{c} & (\text{\textquoteleft lak}) & \text{si} & \text{si} & \text{si} & \text{!} \\
\hline
\end{array}
$$

Although NonFinH, formulated as in (33), directly reflects the phonetic motivation behind it, a constraint banning word-final feet altogether is required for a language like Macedonian, where the main stress invariably is on the antepenultimate syllable (‘type 4 dialects’; cf. Baerman 1999), as illustrated in (38). By ranking NonFinF high, the final two syllables are unavailable for stress.

Q121 Cavineña belongs with Chama and Tacana to the Tacanan language family. Cavineña does not have monosyllabic content words. Chama/Tacana words like $\text{di} \text{ ‘mosquito’ and do} \text{ ‘howler monkey’ are [\textquoteleft \textquoteleft di?i]}$ and [‘do?o)] in Cavineña. Historically, this difference could have arisen in either of two ways. Proto-Tacanan may have been like Cavineña and word-final vowel deletion occurred in Chama/Tacana. Alternatively, Proto-Tacanan was like Chama/ Tacana and a vowel was inserted in Cavineña. Which scenario is the more plausible, considering the data below?

<table>
<thead>
<tr>
<th>Cavineña</th>
<th>Tacana</th>
<th>Chama</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘bi.na’</td>
<td>‘bi.na’</td>
<td>‘bi.ña’</td>
</tr>
<tr>
<td>‘kwi.ha’</td>
<td>‘kwi.da’</td>
<td>‘kwi.sa’</td>
</tr>
<tr>
<td>‘e.do.’</td>
<td>‘e.do. me.so’</td>
<td>‘e.do.’xo.ho’</td>
</tr>
</tbody>
</table>

‘bat’
‘thorn’
‘inside’
Dutch is a quantity-sensitive language which categorically rejects adjacent stressed syllables. Disyllabic words with main stress on the final syllable have a stray initial syllable, as in [ˈriːvər] ‘river’, [ˈkan ˈtoːr] ‘office’, whose duration and voice quality are significantly reduced. Disyllables with main stress on the first syllable include the final syllable as the weak branch of a trochee, as in [ˈpinda] ‘peanut’, [ˈdɔktɔr] ‘holder of doctorate’ (Gussenhoven 1993, 2009). Since the ban on adjacent stresses is blind to the segmental composition of stressless syllables, something needs to be done to circumvent the effect of WSP and CODA-μ, which generally demand that closed syllables are stressed. NoClash (39) serves this purpose by directly targeting the problem, as shown in (41). It takes the location of the main stress for granted. The fact that initial but not final syllables are stray is here attributed to the trochaic nature of the language, but the difference also follows from the right-alignment of trochees.

(39) NoClash: Foot heads are not adjacent.

<table>
<thead>
<tr>
<th></th>
<th>TROCHEE</th>
<th>NoClash</th>
<th>WSP</th>
<th>PARSE-σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>('.dsk.ₜər')</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>('.dsk).ₜər'</td>
<td></td>
<td>*</td>
<td>*!</td>
</tr>
<tr>
<td>c</td>
<td>('.dsk)(.ₜər)'</td>
<td>*!</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(40) dɔk.ₜər Trochee: NoClash WSP PARSE-σ
    a ('.dsk.ₜər') * *
    b ('.dsk).ₜər' * *!
    c ('.dsk)(.ₜər)' *!

(41) kɑn.ˈtoːr Trochee: NoClash WSP PARSE-σ
    a kɑn.(ˈtoːr) * *
    b (ˈkɑn).(ˈtoːr) *!
    c (kɑn.ˈtoːr) *!

11.5 CONCLUSION

Stress systems represent an exciting area of research. In this chapter, we have introduced the concepts that have played a role in typological discussions of word prominence. We drew a distinction between languages with and without word prosodic representations. In the second group, which may be small, no representation beyond a syllabified string of consonants and vowels is present in words, whether lexically or postlexically (Malay). Within the first group, we distinguished between languages with and without phonetic stress. We also observed that tones are present in the lexical or postlexical representation regardless of the presence of phonetic stress. In languages with phonetic stress, the distribution of tones is frequently at least in part determined by the metrical structure, for instance because stressed syllables of accented words have postlexical tone (English, Swedish) or because a lexical tone contrast is only available on the stressed syllable, with possible restrictions on its location (Ma’ya, Swedish). Some languages without phonetic stress have tone distributions that resemble the distributions found for stress systems, in which case the tones may have been described as stress (Nubi, Persian). Tones may be present on or be assigned to words in the lexicon, in which case the language is referred to as a ‘tone language’, or postlexically. Postlexical tones are usually referred to as intonational tones.
Stress was characterized as a property of syllables, not of moras or feet, and as always having a default location. Stresses are properly to be seen as feet, rhythmically motivated constituents potentially enclosing more than one syllable, which we assume explains the obligatory nature of stress. The implication that stress systems are governed by generalizations, either applying unexceptionally or describing the default structures, has attracted a great deal of theoretical work. Section 11.3 set out the main variables in those generalizations on the basis of work by Hayes (1981, 1987, 1995), with ample exemplification. Since stress has been a much-debated as well as relatively successful topic in Optimality Theory, section 11.4 presented the most general constraints which have been used in descriptions of stress systems. These descriptions had a double aim. They not only displayed a wide array of stress systems but also showed how Optimality Theory has been applied to varied but coherent sets of data.
12.1 INTRODUCTION

The phonological structure of languages extends beyond the syllable and the foot. There is a hierarchy of phonological constituents, so that lower, and thus typically smaller, constituents are contained within higher, and thus typically larger, ones. For instance, the foot is contained within, or dominated by, the phonological word, and phonological words are grouped into phonological phrases, etc. Instead of ‘phonological constituent’, the term ‘prosodic constituent’ is often used, particularly for the higher constituents, like the phonological phrase, the intonational phrase and the phonological utterance. The entire structure above the syllable is often referred to as the ‘Prosodic Hierarchy’. This line of research began with Selkirk (1978) and was consolidated by Nespor and Vogel (1986), Hayes (1989b) and, for the phonological word, Booij (1985), among others.

In this chapter, we discuss and illustrate the phonological utterance (U), the intonational phrase (IP), the phonological phrase (φ) and the phonological word (ω). In (1), an example of a sentence is given that has been parsed into these constituents. The next two lower constituents are the foot (chapter 11) and the syllable (chapter 9). Notice, for instance, that weak forms, the function words were and to, are not separate ωs, but are included with the following form in the same ω. In fact, they aren’t even feet, since they have no stress. Before discussing the prosodic constituents in more detail, we first deal with three general questions. First, how do prosodic constituents manifest themselves? Second, what is the general structure of the prosodic hierarchy? And, third, what determines the prosodic structure of specific sentences?

12.2 GENERALIZATIONS INVOLVING PROSODIC CONSTITUENTS

1 Boundary strength. There is a greater degree of articulatory integration in lower constituents than in higher ones. For instance, within an English
syllable like [laemp] as in Lampton, the articulatory gestures occur relatively quickly after each other, but within higher-ranking constituents the sequence of movements in [laemp] will be less tightly coordinated, as in A fu[ll amp]utation or A [lamb p]assed by. That is, prosodic breaks between higher constituents are stronger than those between lower constituents. Boundary strength will reveal itself in a number of phonetic measures. The extent to which the articulation of a vowel is influenced by that of a vowel or consonant in a neighbouring syllable will depend on the strength of the boundary between the syllables containing the vowels. In general, coarticulation will cause a schwa to be closer in [apist] than in [apist], in anticipation of the tongue position of the following vowel, but this effect will be stronger if the two vowels occur within an ω, as in [ap]peal, than when they occur in different ωs, as in Emm[s] Peel (cf. Cho 2004). Also, the duration of the last syllable of a prosodic constituent is typically longer as the lengthening is stronger (or ‘higher’, thinking hierarchically), which is known as final lengthening, also known as preboundary lengthening. Equally, the initial segments of a constituent are more clearly pronounced as the boundary is stronger, which is known as initial strengthening (Keating et al. 2004).

Boundary tones. Higher constituents are often characterized by intona­tional boundary tones. When that happens, the boundary is particularly easy to hear. Examples of this will be given when we discuss the φ and the IP.

Postlexical phonological processes. Prosodic constituents frequently determine the distribution of segments and the application of phonological processes, as illustrated for the syllable and the foot in chapters 9 and 11. The ways in which reference is made to phonological constituents have been classed into three types (Selkirk 1980).

a Domain limit constraints. Reference is made to the left or right edge of a constituent. In many languages, ωs must end in a consonant, as in Tagalog for instance, although syllables occur freely without a coda word-internally.

b Domain span constraints. The context and the focus of some generalization must occur with a single constituent of some rank. For instance, as we will see below, Italian has a rule of s-voicing, which causes [z], never [s], to appear between vowels contained within an ω.

c Domain juncture constraints. The context of the generalization may include the left and right edges of adjacent constituents of some rank, provided this boundary occurs within some higher constituent. For instance, as we will see, [s] is optionally voiced to [z] in Dutch if it occurs finally in the ω and the next ω begins with a vowel, provided this boundary falls within an IP.

Phonology-sensitive syntax. Lastly, it has been observed that syntactic and morphological rules may be sensitive to the size of constituents (Zec and Inkelas 1990). English has Heavy NP Shift, which allows the movement of the object in (3b) to clause-final position. However, a condition on the rule is that the object NP must consist of more than one φ. For this reason, (3a)
cannot move its NP. By contrast, (3b) shows that either order is fine if the object NP contains more than one φ.

(2) \[ V \longrightarrow \text{NP} \rightarrow \text{NP} \]

(3) a He gave the book to her aunt (* . . . to her aunt the book).
   b He gave to her aunt the book about Mozart (or: the book about Mozart to her aunt).

12.3 THE STRICT LAYER HYPOTHESIS

What does the prosodic hierarchy look like? A common view is that the constituents obey the Strict Layer Hypothesis, which is perfectly obeyed by representation (4).

(4) \[ ( \text{IP} ) \longrightarrow ( \phi ) \longrightarrow ( \omega ) \]

There have been a number of formulations of this principle (Selkirk 1984; Ladd 2008: 291), which can be given as (5).

(5) **Strict Layer Hypothesis** (SLH): A prosodic constituent of rank n is immediately dominated by a single constituent of rank n + 1.

An \( \omega \), for example, will be dominated by \( \phi \). If it isn’t, as in (6a), the structure is non-exhaustively parsed, in this case into constituents of rank 2. As it happens, the seriousness of violations of SLH varies considerably. First, non-exhaustive parsing of syllables into feet frequently occurs when word-peripheral syllables are left unfooted (see section 11.3.1), an uncontroversial case of an SLH violation. Thus, segment A in (6a) may begin a syllable without at the same time beginning a foot. Second, if an \( \omega \) were to be dominated by two \( \phi \)s, we would create the incoherent occurrence of a \( \phi \)-boundary within an \( \omega \). This type of violation of the SLH, shown for the boundary between A and B in (6b) and known as ‘improper bracketing’, is ruled out completely. Third, if a constituent were to dominate a constituent of the same rank, we would have recursiveness. A structure is recursive if some constituent appears within a constituent of the same or a lower rank. In (6c), for instance, two constituents of rank 1 are contained within a constituent of the same rank. There have been various claims that same-rank recursiveness in prosodic phrasing must in fact sometimes be recognized.

(6) a \[ ( ) \rightarrow ( )3 \]
   b \[ ( )2 \rightarrow ( ) \rightarrow ( )2 \]
   c \[ ( ) \rightarrow ( )1 \rightarrow ( A \ A \ B ) \rightarrow ( )1 \rightarrow ( )1 \]
As observed in section 1.4, recursiveness in the morphosyntactic structure is one of the hallmarks of human language. It explains why sentence length is infinite, as in *This is the cat that caught the rat that stole the cheese that lay on the table that . . .*, where every NP except *This* has the structure \([\ldots[\ldots]_{\text{NP}}\), as in \([\text{the cat} [\text{the cat stole the cheese}]_{\text{NP}}\), in which the S has an NP which takes the form \([\ldots[\ldots]_{\text{NP}}\), and so on, *ad infinitum*. Similarly, sentences may appear as premodifiers within NPs, as in *an I-couldn't-care-less attitude*. In phonological structure, such spectacular cases of recursiveness are unknown. It will be safe to say, for instance, that we will never come across a language which places a \(\omega\) in the nucleus of a syllable.

### 12.4 FACTORS DETERMINING PROSODIC PHRASING

What determines where these prosodic constituents begin and end? Not surprisingly, an important factor is the morphosyntactic structure. It would, to give an extreme example, be unexpected for the main break in (7a) to occur between *to* and *arrive*, with smaller breaks as indicated. Rather, we would expect something more like (7b). In fact, in (7b), the smaller domains are \(\varphi\)s, and the larger ones IPs. As will be clear, they correspond with syntactic phrases, unlike the bracketed parts of (7a). By somehow marking off the meaningful constituents in the pronunciation, parsing of the expression will be easier for the listener.

\[
\begin{align*}
\text{(7) } & \frac{a}{\{\text{(The first)} \text{ (train to)}\}} \frac{b}{\{\text{(arrive is the)} \text{ (one from Paris)}\}} \\
\text{(7a)} & \{\text{(arrive is the)} \text{ (one from Paris)}\} \\
\text{(7b)} & \{\text{arrive is the)} \text{ (one from Paris)}\}
\end{align*}
\]

However, a secondary role is played by constituent length. Since morphosyntactic constituents of a given rank may vary hugely in length, a one-to-one correspondence between phonological and morphosyntactic constituents would put unreasonable demands on speakers. For instance, the large NP after *see* in (8a) is syntactically equivalent to the word *her* in (8b). It would be quite a strain on the speaker to produce a \(\varphi\) that runs all the way from the *old* to the *road*, while it would equally be awkward to produce two in quick succession for *I can see* and *her*. In (8a) there is too much phonological structure and in (8b) too little for a comfortably rhythmic occurrence of prosodic breaks. Generally, there would appear to be a tendency for constituents to consist of two lower constituents (Selkirk 2000).

\[
\begin{align*}
\text{(8) } & \frac{a}{I \text{ can see the old customs office at the end of the bend in the road.}} \\
\text{b } & I \text{ can see her.}
\end{align*}
\]

Not surprisingly, a \(\varphi\) tends to be produced for each of the NP-internal phrases in (8a): *the old customs office*, *(at the end)*, *(of the bend)* and *(in the road)*. And in (8b) the NP *her* will be incorporated with the preceding *see* into a single \(\omega\), pronounced [siːə], to rhyme with *Maria*. This incorporation of phonologically weak words into adjacent words is called *cliticization*, and *her* here is a clitic that attaches to the host *see*, instead of forming a \(\varphi\), which it would have been entitled to if it had been
a noun. Cliticization of phonologically weak words inside a $\omega$ with a regular word is a specific case of what is called restructuring, the incorporation of phonologically light structures with adjacent words into a phonological constituent of the rank which it would otherwise have had to itself.

In (8a), the correspondence between the syntactic and the phonological structure is still in a sense one to one, since inside the large NP, the ‘maximal projection’, there are smaller phrases, PPs and an NP, or XPs for short, which correspond to $\phi$s. However, also when the syntax is identical, different prosodic structures may be imposed on the grounds of length. For instance, an IP boundary is more likely after the subject NP in Hippopotamuses like to swim in the river than in John likes to swim in the river. Languages will vary in their preference for reflecting the morphosyntactic structure in the phonology at the expense of an even distribution of phonological constituents.

A third factor is the information structure of the sentence. If someone answers the question When was Mozart born? by saying In January 1756, all of the information expressed is new to the speaker who posed the question. However, if the same expression was said in response to Was Mozart born in January 1756 or in February 1756? only January would be the new information. These different focus constituents are indicated in (9a) and (9b), respectively. Languages that encode such differences in information structure, or focus structure, may do so in different ways. Frequently, this is done through contrasts in prosodic phrasing. We will see an example of this in the discussion of the $\phi$ in Bengali. Other languages, like English, use pitch accents for this purpose (section 10.6).

(9)  a  In [January 1756]$_{FOC}$  
      b  In [January]$_{FOC}$ 1756

Alignment constraints will take care of the coincidence of morphosyntactic constituents and prosodic constituents. Selkirk’s (2000) ALIGNXP, for instance, given in (10), can be used to describe the phonological phrasing of the Basque sentence in (11a). It explains why (11b) is ungrammatical, because the right edge of the XP meaning ‘to Amaia’s grandmother’ does not end a $\phi$. The coincidence of prosodic constituents and focus constituents can likewise be accounted for by means of alignment constraints.

(10)  ALIGNXP: Align the right edge of an XP with the right edge of a $\phi$.
(11)  a  $\phi$(Amaien amunari)$_{i_p}$ (liburua)$_{i_o}$ (emon dotzo)$_{i}$  
       Amaia-gen grandmother-dat book give aux  
       She gave the book to Amaia’s grandmother  
      b  *$\phi$(Amaien amunari liburua)$_{i_p}$ (emon dotzo)$_{i}$

An example of a constraint that considers phonological length is Binary (Elordieta 1997), a specific form of the general phenomenon that constituents mustn't be too short or too long. In fact, syntactically well-behaved (13b) is ungrammatical, because (12) outranks (10).
(12) **Binary:** The first $\phi$ of the sentence must contain minimally two $\omega$s.

(13) a $(\text{Amaiari amumen liburua})_{\text{\epsilon}}(\text{emon dotzo})_{\text{\epsilon}}$

    Amaia-dat grandmother-gen book give aux

    She gave grandmother’s book to Amaia

b $^{\text{\epsilon}}(\text{Amaiari})_{\text{\epsilon}}(\text{amumen liburua})_{\text{\epsilon}}(\text{emon dotzo})_{\text{\epsilon}}$

---

**Q122** If the syntactic structure of (13a) were to be used as a response to *Did she give Joseba’s book to Amaia?*, the focus constituent would be *amumen*, the ‘new’ information. The left edge of this kind of focus constituent must be aligned with a $\phi$-boundary, as expressed in (1).

(1) **AlignFOC:** Align the left edge of a FOC-constituent with the left edge of a $\phi$.

The prosodic structure of the reply, in translation ‘She gave GRANDMOTHER’S book to Amaia’, is the one given as ungrammatical in (13b) of the text.

1 How would you account for the fact that the first $\phi$ consists of a single $\omega$?
2 What would the prosodic structure be of the equivalent of ‘She gave grandmother’s BOOK to Amaia’?

---

We now turn briefly to the four prosodic constituents. These are not the only prosodic constituents that are discussed in the literature. For instance, an ‘accentual phrase’ and an ‘intermediate phrase’ are often referred to, both of which would rank above the $\omega$ and below the IP. Their position relative to the $\phi$ may vary across descriptions.

### 12.5 The Prosodic Constituents

#### 12.5.1 The Phonological Utterance

Nespor and Vogel (1986) illustrate the domain span effect of the U, on $r$-linking in the standard variety of English spoken in England, RP (Received Pronunciation). Like many other varieties, RP disallows the [–cons] consonants [h j w r] in the coda. Morphemes that end in nonhigh vowels ([ə ɛ ɔ ə ː]) as in *villa, idea, fair, paw, car, stir* are followed by [r], if the next morpheme begins with an onsetless syllable, as illustrated in (14). In (14a,b), $r$-linking is seen to apply within the word and across words. However, the upper limit is the U: while it can apply across
two sentences addressed to the same listener and not separated by a pause, it cannot
apply across two sentences addressed to different listeners, even if they are spoken
without an intervening pause. The examples (14c,d) illustrate that the U is not nec-
essarily isomorphic with a single syntactic sentence, but that there is nevertheless
an upper limit to what can be accommodated within the same U.

(14) a  stɜːr  (Hi Sheila! [r] Everything all right?)
b  fεər  (A fair idea)
c  ˈʃiːlə  (Hi Sheila! [r] Everything all right?)
d  ˈpiːtə  (Hi Peter!)

12.5.2 The intonational phrase

The IP, also abbreviated as ι, tends to correspond to the root sentence, i.e. a single [NP VP] structure without extrapositions or interruptions. Selkirk (1978)
gives (15b), in which the extraposition in Pakistan and the restrictive relative clause which is a weekday have been assigned to separate IPs, leaving the root
sentence, which would otherwise be a single IP (15b), to be divided over two IPs.
However, as was the case with U, the division of speech into IPs is not purely
syntactically driven. In particular, when the subject is longer than a single lexical
word there will tend to be an IP boundary between the subject NP and the VP, as shown in (15c).

(15) a  (In Pakistan), (Tuesday), (which is a weekday), (is a holiday),
b  (Tuesday is a holiday),
c  (The second Tuesday of every month), (is a holiday),

The domain span effect of the IP can be illustrated with a rhythmic accentuation
affecting certain adverbia ls in Dutch, like [alteit] altijd ‘always’, which can appear
in a variety of sentential positions in the IP. They are accented on the final syllable
when no other pitch accented word follows in the same IP, as illustrated in (16a,b). The retraction of the accent to the first syllable occurs when they are followed by an
accented word within the IP, as illustrated in (16c,d).

(16) a  (Naar de WAterstanden luistert ze altijd)
     to the water level reports listens she always
     ‘The water level reports she will always listen to’
b  (Waar ze altijd naar luistert), (zijn de WAterstanden),
c  (ALTijd luistert ze naar de WAterstanden),
d  (Ze luistert ALTijd naar de WAterstanden),

As illustrated in chapter 10, the IP is bounded by intonational boundary tones in
many languages, a domain limit phenomenon. English non-final IPs are frequently
closed by H% after a H*L pitch accent, which causes the final syllable of incident
in (17a) to have high pitch. In British English, the same H*L H% pattern may be used
for questions, as shown in (17b).
12.5.3 The phonological phrase

Although languages vary in the details of the correspondence, the $\varphi$ tends to correspond to the syntactic phrase. Hayes (1989b) shows that the $\varphi$ defines the domain of the English rhythm rule. In (18a) Chinese is an adjective inside the NP the Chinese dishes, while in (18b) the Chinese is an NP, the indirect object of gives. An adjustment of the accentuation pattern from CHINESE to CHInese under the influence of the following accent on DISHes takes place in (18a), but not in (18b). Similarly, the German rhythm rule is sensitive to the German $\varphi$ in (19a), the headless object NP das hell-blaue forms a $\varphi$ by itself, while in (19b) it forms part of the NP, and hence the $\varphi$, das hell-blaue Bild.

(18) a On Tuesdays, he gives (the CHInese DISHes)

b On Tuesdays, he gives (the CHINESE)

(19) a Ich fand (das HELL-BLAAe) (SCHÖN)

‘I found the light-blue one beautiful’

b Ich fand (das HELL-blaue BILD)

‘I found the light-blue picture’

In Bengali, the $\varphi$ is phonologically marked by a final boundary tone $H_\varphi$ if it contains an intonational pitch accent $L^*$, as shown in (20a). Because the right-hand boundary of the $\varphi$ is sensitive to the focus of the sentence in Bengali, (20a) contrasts with (20b), whose focus is confined to the first constituent of the compound word for ‘fishhead’. Bengali also has IP-final boundary tones, like $L_\varphi$. The Bengali $\varphi$ equally defines the domain of the rule of r-deletion and a regressive voicing process (Hayes and Lahiri 1991; see Q48).

(20) a tumi $\varphi$ (kon mat$^b$er-mat$^a$) $\varphi$ (ranna-korle) $\varphi$

‘Which FISHHEAD did you cook?’

b tumi $\varphi$ (kon mat$^b$er$^b$) $\varphi$ (mat$^b$ a) $\varphi$ (ranna-korle) $\varphi$

‘The head of which FISH did you cook?’
Again, syntactic constituency does not provide the only relevant information for deriving \( \varphi \)-structure. Constituents to the right of the NP head in English, such as the PP of ancient China in the NP the CHINESE of ANCient CHina, form their own \( \varphi \), as shown by the absence of an effect of the rhythm rule on the word Chinese. However, when the postposed PP does not contain a lexical head, like the PP on him in the NP that report on him, a single \( \varphi \) is formed containing the whole NP. Recall from section 12.2 that Zec and Inkelas (1990) pointed out that the syntactic rule of heavy NP shift, which places an object NP in sentence-final position, only applies if the NP is composed of at least two \( \varphi \)s. This is shown in (21), taken from Inkelas (1989). In the ungrammatical b-example, the object is only a single \( \varphi \), thanks to the prosodically restructured PP on him.

(21) a Mary gave to Susan (that report)\(_{\varphi}\) (on Dukakis)\(_{\varphi}\)  
    b *Mary gave to Susan\(_{\varphi}\) (that report on him)\(_{\varphi}\)

The examples in (22a,b) (Nespor and Vogel 1986) illustrate another way in which nonsyntactic information is relevant to \( \varphi \)-formation in (American) English, as revealed by the behaviour of the rhythm rule. In (22a), [ˌrɪˈprəˈdjuːs] reproduce has the main stress on the last syllable, which pronunciation is as expected when it is the last word in the \( \varphi \). However, in (22b) it has undergone the rhythm rule, a pronunciation that requires that it should be followed by another accented word in the same \( \varphi \). This is explained by the assumption that if the Adverbial Phrase consists of a single \( \omega \), it can optionally be included in the \( \varphi \) on its left, instead of forming its own \( \varphi \).

(22) a Rabbits\(_{\varphi}\) (REproDUCE\(_{\varphi}\)) (QUICKly and DILligently)\(_{\varphi}\)  
    b Rabbits\(_{\varphi}\) (REproduce QUICKly)\(_{\varphi}\)

Q123 What is the \( \varphi \)-structure of sentences (1) and (2)? Why is (2) ungrammatical?

(1) I was explaining to the students the problem of the double negation in Middle English.  
(2) *I was explaining to the students the problem.

Q124 The English rhythm rule can apply in (1) but not in (2). How would you account for this difference?

(1) This is REPresented in SIX. (i.e. ‘in (6)’)  
(2) This is REpreSENted in SIX A. (i.e. ‘in (6a)’)
12.5.4 The phonological word

Frequently, when a phenomenon is said to be word-based, it is in fact confined to the domain of the $\omega$ (also known as the prosodic word). Crucially, $\omega$ does not correspond in a one-to-one fashion to the morphological word. For instance, while compounds represent ‘words’ in the sense of morphological categories like Noun, Verb or Adjective, in many languages each of the constituent parts forms a phonological domain for (consonant or vowel) harmony, word stress and syllabification. Thus, vowel harmony in Turkish (cf. section 6.4.2) is confined to the constituents of the compound, as shown in (23), where the vowels in the first constituent are $[-\text{back}]$ and the vowels in the second are all $[+\text{back}]$. In German, the Maximum Onset Principle (section 9.2) does not apply across the internal boundary in a compound, as shown in (24). In Greek, vowel hiatus between the members of the compound is tolerated, but not within non-compound words, while each of the constituents has its own word stress, exactly as if they formed an NP like [me ‘yalia erya’sia] ‘big works’, as shown in (25) (Nespor and Vogel 1986; Nespor 1998).

\[(23)\]  
\begin{center}
\begin{tabular}{ll}
a & kopek balugu (*kopek beligi) \quad & ‘shark (lit. dog fish)’ \\
b & el jazusu (*el jezisi) \quad & ‘handwriting’ \\
\end{tabular}
\end{center}

\[(24)\]  
\begin{center}
\begin{tabular}{ll}
a & mont.a rt (*monta:rt) \quad & ‘dialect (lit. mouth type)’ \\
b & ais.loef (*ais:leff) \quad & ‘ice spoon’ \\
\end{tabular}
\end{center}

\[(25)\]  
\begin{center}
\begin{tabular}{ll}
a & o’moa े े े sias (*oma:erya’sias) \quad & ‘work team’ \\
b & ‘zoni asfa’ lias (*zonasfa’lias) \quad & ‘safety zone’ \\
\end{tabular}
\end{center}

Italian s-voicing, shown in (26), provides an example of an $\omega$-domain span rule. It voices nongeminate [s] between vowels, as illustrated in (27) (Nespor and Vogel 1986). The rule applies in a simplex word in (27a) and in a suffixed word in (27b), but does not apply across a prefix and its base (27c) or a combination of a stem and a word (27d), even though all four items in (27) are single morphological words. The generalization that brings this disparate group of morphosyntactic constituents under one heading is the $\omega$.

\[(26)\]  
\begin{center}
\begin{tabular}{|c|c|}
\hline
V & C V \\
\hline
\end{tabular}
\end{center}

\[(27)\]  
\begin{center}
\begin{tabular}{ll}
a & ’kaza \quad & ‘house’ \\
b & ka’zina \quad & ‘house+DIM’ \\
c & a so t’fale \quad & *azo t’fale \quad & ‘a-social’ \\
d & ‘fiло so’vjetiko \quad & *’filozo’vjetiko \quad & ‘pro-Soviet’ \\
\end{tabular}
\end{center}

A constituency effect of the Dutch $\omega$ was noted by Booij (1985). Dutch coordinated NPs allow deletion of identical parts in the coordinated constituents. Schematically, the situation can be represented as $AB$ and $CB$, where the slash marks the deleted item. Crucially, the deleted portion in the left-hand coordinate is not a morphological (or syntactic) constituent. This is illustrated in (28). The deleted $B$ corresponds to the head noun of an NP in (28a), to a noun within a compound

\[(28)\]  
\begin{center}
\begin{tabular}{ll}
a & kaza b \quad & ‘shark (lit. dog fish)’ \\
b & el jazusu (*el jezisi) \quad & ‘handwriting’ \\
c & a so t’fale \quad & *azo t’fale \quad & ‘a-social’ \\
d & ‘fiло so’vjetiko \quad & *’filozo’vjetiko \quad & ‘pro-Soviet’ \\
\end{tabular}
\end{center}
noun in (28b), to a verbal stem in (28c) and to an adjectival suffix in (28d). While the deleted portions are quite heterogeneous when viewed from a morphosyntactic perspective, their common phonological characteristic is that they form separate syllabification domains, i.e. $\omega$s. Dutch does not generally syllabify across words, while all prefixes and certain suffixes like -schap and -achtig do not syllabify together with the base they are attached to, forming separate $\omega$s.

The assumption that the deleted portion should minimally be an $\omega$ is supported by the impossibility of deleting suffixes that do syllabify with the base. Dutch has two adjective-forming suffixes meaning 'like'. The suffix -achtig [ɑχt] is like -schap in (28d), and forms its own $\omega$, but the suffix -ig [əχ] is incorporated into the $\omega$ of its base. Accordingly, deletion of -achtig is possible in (29a), while in (29b) no deletion is possible.

Q125 In Dutch, prevocalic [s] is often voiced to [z] after voiced segments, as shown in (1). However, no voicing is possible in the examples in (2). What determines when [s] may be voiced?

   [ˈvɑs]$_{Aux}$  [ˈɑ:rðɔχ]$_{Adj}$  [z]  ‘was friendly’
   [ˈeis]$_N$  [ɑχtɔχ]$_{Adj}$  [z]  ‘ice-like’
2. [ˈheis]$_{N}$  [ɔn]$_{Y}$  *[z]  ‘hoist+INF’
   [ˈkɑns]$_N$  [ɔn]$_{N}$  *[z]  ‘chance+PL’
   [ˈmɑsa:]$_N$  *[z]  ‘mass’
12.6 DERIVING PROSODIC CONSTITUENTS

With the \( \omega \) we have come to the lowest prosodic constituent which can somehow be related to morphosyntactic constituency. What the \( \omega \) and higher constituents have in common is that at least part of their formation is dependent on the morphosyntactic structure of the language. The question of how the relation between the two kinds of constituency is to be expressed has received different answers in the literature. In Nespur and Vogel (1986), the relation is based on a variety of morphosyntactic properties. A typical statement of such a relation might be ‘Include the head of the syntactic constituent \( S \), together with all the prosodic constituents of rank \( C \) on its non-recursive side, in Prosodic Constituent of rank \( C+1 \).’ (The non-recursive side is the left side in right-branching structures, and vice versa.) Selkirk (1986), with reference to a proposal in an earlier version of Chen (1987), suggested that the unifying element in the relation between prosodic and morphosyntactic constituency is reference to edges. This approach lay at the basis of the development of alignment constraints in Optimality Theory (chapter 4).

We illustrate edge alignment with the Dutch \( \omega \). It can be derived with the help of morphological information only (cf. Booij 1977: 103; van der Hulst 1984: 85). As is the case in Italian, suffixes are syllabified with their base, but prefixes never are. This suggests that the \( \omega \) co-begins with the beginning of the morphological category ‘word’, i.e. any stem or derived word, as expressed in (30). This excludes suffixes and, as we will see in section 12.6.1, certain function words, like pronouns and prepositions. Constraint (30) correctly predicts that prefixes and constituents of the compound form individual \( \omega \).

(30) Align(morphological word, \( \omega \), Left): The left edge of a morphological word aligns with the left edge of \( \omega \).

In (31a), the prefix begins a lexical category (the complex word), and so does the base. (The end of a non-final \( \omega \) is of course defined by the beginning of the next.) Similarly, (31b) illustrates how separate domains are created for the constituents of compounds: each of them begins a lexical category, while the first, additionally but redundantly, begins the compound, another lexical category. In (31c), suffixes are included in the \( \omega \) on the left, because suffixes do not begin lexical stems.

(31) **Morphology**

<table>
<thead>
<tr>
<th>A</th>
<th>[snt [ei.] [on]]</th>
<th>( \omega )-structure</th>
<th>Syllabification</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[snt [ei.] [on]]</td>
<td>( \omega )-structure</td>
<td>'dispossess'</td>
</tr>
<tr>
<td>b</td>
<td>[rein [a:k] [œyl]]</td>
<td>'Rhine barge'</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>[te:kon [na]]</td>
<td>'barn owl'</td>
<td></td>
</tr>
</tbody>
</table>

As pointed out by Inkelas (1989) and Booij (1996), prosodic structure, like any other aspect of the phonological representation of words or morphemes, can be included in underlying representations. As we saw in (28d) and (29) above, many
Dutch full-vowelled suffixes are not included in the $\omega$ on their left, like the nominalizing suffix -\textit{schap} and the adjectival suffix -\textit{achtig}, which do not syllabify with their base. In the lexicon, these suffixes will therefore be listed as $\omega$s.

Selkirk (2011) re-evaluates the pervasive tendency for syntactically defined constituents to correspond to prosodic constituents like $\omega$, $\phi$ and $\iota$. Instead of requiring coincidence of either left or right edges, Match Theory claims that the correspondence between, say, a syntactic phrase and $\phi$ is direct, with edges on both sides coinciding. For some recent work on this position, see Selkirk and Lee (2015).

### 12.6.1 Clitics

Pronouns, auxiliary verbs, conjunctions and the like cannot be given $\omega$ status by (30) in Dutch, since they do not belong to a major word class. But since they are words, not affixes, they cannot attach to some other item in the lexicon that is a major-class item. Many function words are in fact included in $\omega$s postlexically. If (30) is also valid postlexically, the prediction is that function words in Dutch should behave like suffixes, i.e. be included in the $\omega$s to their left. This is indeed what we find. In (32a) the article [\textit{an}] encliticizes onto the preceding verb form [\textit{rip}] ‘called’. This explains why the article cannot, in natural speech, be pronounced [\textit{ʔan}] in this context. The same goes for the preposition [\textit{in}] and the definite article [\textit{at}] in (32b) (Booij 1996).

\begin{verbatim}
(32)  a  [\textit{rip}] \textit{an} \textit{kat}    \textit{Hij (ri.pan)} \textit{kat}
       ‘he called a cat’
    b  [\textit{za:t}] \textit{in} \textit{prop} \textit{at} \textit{putje} \textit{Het (za:ti.nat) putje}
       ‘it goes into the drain’
\end{verbatim}

Other function words, also those that lack an onset, have a full vowel, like the conjunctions [\textit{of}] ‘or’ and [\textit{en}] ‘and’. These words are at best only variably syllabified with the preceding word, and they will therefore have to be given $\omega$s status in the lexicon, along with the suffixes that form their own $\omega$s, like -\textit{schap}. It has also been noted that certain function words and affixes behave neither as elements that are included in the same $\omega$s as their host word nor as elements that form an $\omega$ by themselves. Unlike other words, Italian object pronouns have no stress, but they nevertheless maintain their status as $\omega$s, as shown by their phonology at the boundaries with stressed words. For one thing, they do not change the location of the stress on the ‘real’ words to their left, as in [\textit{teˈlefona lo}] ‘Call him!’, whose pre-antepenultimate stress would be ungrammatical if [lo] were part of the same word. For another, \textit{radoppiamento syntattico}, which geminates a word-initial consonant after a stressed vowel, as in [ˈdal ˈli:brei] ‘Give books!’, as in a listing like \textit{Don’t donate clothes. Give computers! Give pens! Give books!}, from [ˈda ˈli:brei], equally applies to pronouns, as in [ˈdal lo] ‘Give it!’ from [ˈda lo]. On the basis of such data, Nespor and Vogel (1986) postulate the Clitic Group as a constituent immediately above the $\omega$, which makes it possible to account for phonological processes that occur between words and such recalcitrant morphemes, but that
fail to occur in other morpheme combinations. Alternative accounts have been presented, however.

12.6.2 The syntactic residue

The U, i, ϕ, ω and, perhaps, the Clitic Group are the prosodic constituents which define the relevant domains of processes that apply above the word level. Prosodic theory thus distinguishes itself from theories that claim that such rules can refer directly to syntactic structure, such as Kaisse (1985). Nevertheless, instances have been found of rules that apply across words which do apparently refer to syntactic categories, as would appear to be the case for French liaison (Hayes 1990; Post 2000). Hayes (1990) proposes that such ‘residual’ syntax-sensitivity should be accounted for in the lexicon. That is, the phonological rules that produce the required forms are in fact lexical rules, and the forms they produce are thus available in the lexicon, ready for insertion into syntactic phrases. For instance, the phonological rule that shortens final long vowels in Hausa verbs is syntactically conditioned: it only applies if a direct object that contains a major-class noun immediately follows. This is illustrated in (33a), which contrasts with (33b), two cases in which the morphosyntactic condition is not met.

(33) a náː káːmà kíː fíː ‘I have caught a fish’
b náː káːmáː ʃí ‘I have caught it’
náː káːmáː wà múː sáː kíː fíː ‘I have caught Musa a fish’

The lexical rule is given in (34); its morphosyntactic conditioning is expressed by the ‘Frame’ given below it. (Because the only category that can occur initially in a VP before an NP is a V, it is not necessary to label the word as a Verb in the rule.) Thus, when a verb is to be inserted in a sentence, the more specific form produced by final vowel shortening is chosen if the morphosyntactic condition applies.

(34) **Final vowel shortening**

\[ V \rightarrow \emptyset / \ldots V \_{Frame} ] \_{Frame 1} \\
Frame 1: [ \_ NP \ldots ]_{VP} \]

The assumption that syntax-dependent rules are in fact lexical rules puts such alternations in a comparable position with phrasal allomorphy of the sort that is seen in the English indefinite article, which is [ən] before vowels but [ə] elsewhere. As the name suggests, phrasal allomorphs are rival phonological forms whose distribution is governed by properties of the surrounding words. There are, however, two differences between these two cases worth mentioning. First, the forms in (34) are generated by a rule, because they involve a whole class of words rather than a single morpheme, and, second, in (34) the conditioning is morphosyntactic rather than purely phonological. Hayes refers to forms like English [ən] and Hausa [káːmáː] as *precompiled*, the idea being that they come ready-made from the lexicon. A prediction of this treatment is that rules like (34) might have exceptions. In fact, in the context of (34), a restricted group of Hausa verbs, e.g. ['sàjáː] ‘bought’, raise their final [əː], in addition to shortening it, as in [náː sàjì əbìnci] ‘I bought food’.
Phonological rules that apply above the level of the word are constrained by prosodic constituents that form a hierarchical structure. These constituents are not directly mappable onto the morphosyntactic structure. While the prosodic hierarchy above the $\omega$ is in large part derivable from the syntactic structure, mismatches do occur, mainly as a result of the incommensurate lengths of the phonological constituents (cf. Selkirk and Lee 2015). A syntactic constituent may be too long to fit into the phonological constituent it usually maps onto, with the result that it is divided over a number of such constituents, or too short, in which case it may be included in a constituent of the same (restructuring) or a lower rank (cliticization).

The structural separation of the morphosyntactic and phonological grammars may appear to be compromised in two ways. First, we have seen that prosodic constituents not only define the domains at or within whose edges phonological adjustments occur, but are arguably also referred to by rules of syntax, as in the case of heavy NP shift. This may indicate that, just as foot structure may be available in the lexicon (cf. chapter 11), so morphosyntactic structure will exist simultaneously with the phonological structure. For the reverse case, phonological rules which appeal to syntactic information, Hayes (1990) provided a solution which upholds the distinction between lexical rules, which can refer to morphological information and which potentially have exceptions, and postlexical rules, which can only refer to phonological representations and which cannot have exceptions. He proposed that phonological rules for which syntactic information is relevant are in fact lexical rules. Such syntax-sensitive rules only apparently apply above the level of the word and actually produce the appropriate alternants in the lexicon. During the construction of the sentence, these precompiled forms are inserted in the specific contexts instead of their rival, more general alternant.

Prosodic structure appears to be more orderly than syntactic structure. The SLH, which forbids improper bracketing, recursivity and non-exhaustiveness, constrains the prosodic structure from the $\omega$ onwards reasonably successfully, since deviations appear to be limited. Phonological constituents below the $\omega$, which are not derivable from the morphosyntactic structure, deviate from the SLH in limited ways. In particular, not all syllables need to be included in foot structure (stray syllables, section 11.3.1). We may safely predict that prosodic phonology will continue to be an exciting field of research for some time yet.

**NOTE**

1 The segment is known as ‘linking $r$’ if there is an $<r>$ in the spelling and ‘intrusive linking $r$’ if there is not. In the latter case, speakers may feel that the pronunciation if [r] is ‘incorrect’. 

13.1 INTRODUCTION

In earlier chapters, we have seen how phonological representations have changed from linear strings, as in the early SPE model, to non-linear ones with autonomous and parallel strings of vowels, consonants and tones as well as a hierarchical organization of constituents encompassing these segments. We have also seen two ways in which adjustments in these representations as they are incorporated in ‘later’ morphosyntactic structures have been described. One way was to allow rules to change representations, and the other was a competition among possible output representations as decided by a series of ranked criteria, the constraints of Optimality Theory (OT).

The last two chapters of this book return to each of these broad topics. Chapter 14 will consider the question whether segments have internal structure. In the SPE model, the segment consisted of a list of distinctive features, without any internal ordering. However, a strong case has been made for the assumption that segments are trees whose branches represent specific groups of features, like those defining the place of articulation of a segment. Chapter 13 will point out that while OT has booked some striking results that are unattainable in a rule-based framework, like the avoidance of the duplication problem, there are things a rule-based analysis can do that seem unattainable in OT.

We will start by discussing a number of issues in rule ordering which have preoccupied many a phonologist over the years. First, are there general principles that predict the order in which rules apply? Are there rule orders that are more natural or more expected than others? The work on these questions has led to the recognition of four types of rule interaction: the feeding, bleeding, counterfeeding and counterbleeding orders. Additionally, we draw attention to the notion of opacity. A surface form may look as if some rule should have applied to it, but didn’t, or it may look as if some rule applied to it, but without apparent need, because the appropriate phonological context isn’t there. Such opaque forms arise as a matter of course from rule ordering, as we will illustrate, but are predicted not to exist by any theory in which all constraints apply to surface forms. Solutions to the opacity problem in OT have come and gone in rapid succession. In this chapter, we discuss two of them. One is in line with the theory of Lexical Phonology described in chapter 8, and accordingly splits the OT grammar up into minimally a lexical and a postlexical component, known as Stratal Phonology. The second involves an enrichment to the evaluation procedure known as ‘chains’, which attempt to
reproduce the creation of intermediate forms, as standardly produced after every rule application, in OT.

### 13.2 EXTRINSIC AND INTRINSIC ORDERING

Is the order in which rules apply predictable from any properties of the rules concerned? If it is, no ordering statement would be necessary: the rule order is said to be intrinsic. If the order is not given by the theory, and an explicit ordering statement of the type ‘Rule X applies before Rule Y’ is necessary, the rule order is extrinsic. The issue of intrinsic rule ordering occupied many phonologists in the 1970s, but the search for the principles that exhaustively govern the order in which rules apply is generally considered to have been unsuccessful (Kenstowicz and Kisseberth 1977).

A principle that has stood the test of time is the ELSEWHERE CONDITION. This is really a principle governing the application of rules in general, and has been invoked in morphology as well as phonology. What it says is that when one rule applies to a subset of the forms that another rule applies to, the general rule is blocked from applying to that subset. A morphological example will make it clear why this is a useful principle. Take the English morphological rules given in (1) and (2). Rule (1) says: ‘Attach the suffix [z] to noun stems in order to form the plural’, and rule (2) says: ‘Attach the suffix [ən] to the noun stem ox in order to form the plural’.

(1) [kn]z]\text{pl}
(2) [ɔks][n ən]\text{pl}

In order to prevent the formation of *oxes, we must either stipulate that (2) applies before (1), or add to (1) the clause ‘except in the case of [ɔks]’. The ELSEWHERE CONDITION makes either move unnecessary: because (2) applies to a subset of the contexts specified by (1), it automatically blocks (1). This principle thus saves us from having to add all sorts of exception clauses to general rules. A Finnish phonological example, from Kiparsky (1973), is given in (3).

<table>
<thead>
<tr>
<th>Underlying</th>
<th>Derived</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>menek</td>
<td>mene</td>
<td>‘go’</td>
</tr>
<tr>
<td>menek alas</td>
<td>mene alas</td>
<td>‘go down’</td>
</tr>
<tr>
<td>menek pois</td>
<td>menep pois</td>
<td>‘go away’</td>
</tr>
<tr>
<td>menek kotiin</td>
<td>menek kotiin</td>
<td>‘go home’</td>
</tr>
</tbody>
</table>

Word-final [k] in Finnish is deleted, unless a consonant follows, in which case the [k] assimilates to that consonant, creating a geminate. The k-ASSIMILATION rule is given in (4). The rule at issue is k-DELETION. Without the ELSEWHERE CONDITION, we would have to state this rule as (5), which explicitly specifies the context ‘when followed either by a vowel or by a pause’, i.e. ‘except when a C follows’.

<table>
<thead>
<tr>
<th>(4) k-ASSIMILATION</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>k # C → 3 2 3</td>
<td>1 2 3</td>
</tr>
</tbody>
</table>
(5) k-deletion \[ k \rightarrow \emptyset \quad /\_\# \begin{cases} V \\ \text{pause} \end{cases} \]

If we were to formulate k-deletion as (6), the elsewhere condition would ensure that it applies only if (4) did not apply. Rule (4) applies to a proper subset of the forms to which (6) is applicable, and therefore applies first, blocking (6) from applying to [menek pois]. Also, (6) will not apply to [k] in [menek kotii], which after all was input to rule (4). In this form, (4) assimilated the ‘original’ [k] of [menek] to the [k] of [kotiin]. Although it applied vacuously, it did apply, thereby blocking (6).

(6) k-deletion \[ k \rightarrow \emptyset \quad /\_\# \]

A widely cited data set used for arguing for extrinsic ordering is based on dialectal variation in Canadian English (Joos 1942), which can be described as resulting from different orderings of the same two rules, flapping and pre-fortis clipping (the terms are from Wells 1982). The first rule, given in (7), causes [t d] to be pronounced as an alveolar tap before reduced syllables, as in [ˈsiri] city. Note that [ɾ] is voiced. The second rule (8) shortens vowels and sonorant consonants preceding voiceless segments. As a result of this rule, the [i:] in [bi:d] bead is longer than that in [bi:t] beat, the [ɛn] in [tenz] tense is longer than the [ɛn] in [tens] tense, and the [i:] in [ˈti:zɪŋ] teasing is longer than that in [ˈliːzɪŋ] leasing. (More adequate statements of these rules are possible, but not necessary for the point at issue.) We will use an ad hoc feature [±long].

(7) flapping \[ [t d] \rightarrow [ɾ / [−\text{cons}]] \begin{cases} V \\ \text{stress} \end{cases} \]

(8) pre-fortis clipping \[ [+\text{voice}] \rightarrow [−\text{long}] / [−\text{voice}] \]

In some dialects, these rules applied in the order flapping – pre-fortis clipping. That is, the words rider (‘someone who rides’) and writer (‘someone who writes’) are homophones, both being [ˈrɛɪər], while ride and write are [ræɪd] and [ræɪt] (where [æi] represents a shortened [ai]). This is shown in (9).

(9) \[
\begin{array}{cccc}
\text{Rule (7)} & \text{ræɪər} & \text{ræɪər} & \text{ræɪd} & \text{ræɪt} \\
\text{Rule (8)} & (n.a.) & (n.a.) & (n.a.) & (n.a.) \\
\text{Output} & \text{ˈræɪər} & \text{ˈræɪər} & \text{ræɪd} & \text{ræɪt} \\
\end{array}
\]

Today, dialects like (9) have apparently become extinct (Kaye 1990). In Canadian English today, these words are pronounced [ræɪər], [ræɪər], [ræɪd] and [ræɪt]. This situation is obtained if we reverse the order of the rules, as shown in (10). But if dialects can differ depending on the order in which two rules apply, as indeed they can, it cannot be the case that rule ordering is predictable (see (26) in section 6.5).
Early attempts to find universal principles governing the order in which rules apply led to a categorization of rule interactions (Kiparsky 1968). The idea was that certain rule orderings were more natural or expected than others, and that phonological change could in part be explained by assuming that in the course of time rules tend to reorder so that they come to have natural orders. While the principles have been abandoned, the terms used to refer to the different types of interaction have acquired wide currency. Four types will be distinguished. In every case two rules are assumed, which will be referred to as rule A and rule B.

13.3.1 Feeding order

If rule A increases the number of forms to which rule B can apply, the order \( A - B \) is a feeding order. The British English data in (11) illustrate preglottalization, given in (12), a rule which glottalizes voiceless plosives in the coda, discussed in section 8.6.3.

(11) Underlying Derived
lok\(t\) lo\(\ddot{\text{k}}\)t looked
k\(\ddot{\text{a}}\)ts k\(\ddot{\text{e}}\)ts cats
h\(\ddot{\text{i}}\)nts h\(\ddot{\text{n}}\)ts hints
k\(\ddot{\text{a}}\)mp\(\ddot{\text{g}}\)raʊnd k\(\ddot{\text{a}}\)mp\(\ddot{\text{\alpha}}\)nd campground

(12) Preglottalization
\[
\begin{align*}
\text{cont} & \rightarrow + \text{constr} / + \text{voice} \quad \text{(# #)} \quad \text{C}
\end{align*}
\]

Now consider the data in (13). In the first and second columns, we see that between a nasal and a voiceless fricative in the same syllable a voiceless plosive is inserted, whose place of articulation is that of the preceding nasal. As shown in the third column, this rule of voiceless stop insertion (14) applies before (12), and thus causes words like prince, which underlyingly end in a nasal consonant followed by a fricative, to be input to rule (12). That is, (14) feeds (12).

(13) Underlying Fortis stop Preglottalization
pr\(\ddot{\text{m}}\)s pr\(\ddot{\text{r}}\)nts pr\(\ddot{\text{m}}\)ts prince
l\(\ddot{\text{e}}\)ŋ\(\ddot{\text{\theta}}\) l\(\ddot{\text{e}}\)ŋ\(\ddot{\text{\kappa}}\)ŋ l\(\ddot{\text{e}}\)ŋ\(\ddot{\text{\kappa}}\)\(\ddot{\text{\theta}}\) length
w\(\ddot{\text{\alpha}}\)\(\ddot{\text{\m}}\)\(\ddot{\text{\tau}}\) w\(\ddot{\text{\alpha}}\)\(\ddot{\text{\m}}\)\(\ddot{\text{\tau}}\)\(\ddot{\text{\theta}}\) warmth

(14) Voiceless stop insertion
\[
\begin{align*}
\emptyset & \rightarrow \begin{aligned}
- \text{cont} & \rightarrow - \text{voice} / + \text{nas} & + \text{cont} & \rightarrow \sigma
\end{aligned}
\end{align*}
\]
13.3.2 Counterfeeding order

The opposite of feeding is counterfeeding. If rule A increases the number of the forms to which rule B can apply, the order B – A is a counterfeeding order. The French feminine adjectival suffix is [ə]. The language has a rule that deletes word-final [ə], which is given as \( \sigma \)-deletion in (15). The language also has final consonant deletion (16), which deletes certain word-final consonants in contexts other than before a vowel or glide, causing [PATIT] to be pronounced [PATI] in the context concerned. Final consonant deletion applies before \( \sigma \)-deletion. If the rules applied in the opposite (feeding) order, \( \sigma \)-deletion would cause the feminine alternant of ‘little’ to be homophonous with the masculine form. In other words, the rules apply in counterfeeding order, as illustrated in (17) for the forms for ‘little nephew’ and ‘little niece’.

(15) \( \sigma \)-deletion \( \sigma \rightarrow \emptyset /\_\# \\
(16) \) final consonant deletion \( C \rightarrow \emptyset /\_\# \ [+\text{cons}] \\
(17) \) Rule (16) \( \sigma \rightarrow \emptyset /\_\# \\
\) Rule (15) \( \sigma \rightarrow \emptyset /\_\# \\
\) patit nav\( \sigma \) patit-\( \sigma \) njes

13.3.3 Bleeding order

If rule A decreases the number of the forms to which rule B can apply, the order A – B is a bleeding order. This type of interaction, in which rule A prevents rule B from applying to particular forms, occurs in the derivation of the English plural. In chapter 4, it was shown that \( i \)-insertion prevents devoicing from applying to a form like [b\( \alpha \)s-z] by separating the final [z] from the stem-final obstruent. In (18), this interaction is shown. That is, \( i \)-insertion bleeds devoicing.

(18) \( i \)-insertion \( (n.a.) \) kis-z \( (n.a.) \) au-z \\\n\) devoicing \( s \ (n.a.) \ (n.a.) \) \\
\) beks kisz aiz

Thus, when two rules A and B are in a counterfeeding order, the application of rule A does not increase the number of forms to which rule B can apply. When they are in a bleeding order, rule A actually decreases the number of forms to which rule B can apply.

13.3.4 Counterbleeding order

The opposite of bleeding is counterbleeding. If rule A decreases the number of the forms to which rule B can apply, the order B – A is a counterbleeding order. If two rules can apply to the underlying form, this order enables both rules actually to do so. In the Kaatsheuvel dialect of Dutch, [\( \sigma \)] is inserted between the noun stem and the diminutive suffix, if the stem ends in a lax vowel followed by a sonorant consonant. Thus, we find it in (19a), but not in (19b), where the stem ends in an
obstruent, or in (19c), where the stem has a tense vowel. The rule of $\alpha$-INSERTION is
given in (20). Here, X is a free variable and represents whatever shape the diminutive
suffix has.

(19) a  snɔr-kə  snɔrəkə  ‘moustache’
    hal-kə  haləkə  ‘hall’
    kum-kə  kuməkə  ‘comb’

b  lɑp-kə  lɑpkə  ‘cloth’

mʏs-kə  mʏskə  ‘sparrow’

c  raːm-kə  raːmkə  ‘window’

(20) $\alpha$-INSERTION $\emptyset \rightarrow \alpha / [-\text{tense}] + \text{cons} + _{\text{X}}\text{DIM}$

Kaatsheuvel also has a rule which inserts [s] between the diminutive suffix [kə] and a stem-final dorsal consonant. Thus, s-INSERTION (21) breaks up a sequence of two dorsal consonants, as in [bakska], from [bak-kə] ‘tray’, [maːxske], from
[maːx-kə] ‘stomach’.

(21) s-INSERTION $\emptyset \rightarrow s/ [+\text{cons} + \text{DORSAL}] + _{\text{kə}}\text{DIM}$

If we want to know how these two rules interact, we need to consider the diminutive
form of a word like [slɑŋ] ‘snake’, which satisfies both rules. Let us suppose,
contrary to fact, that it is *[slɑŋkə]. In order to obtain this form, we would have to
apply $\alpha$-INSERTION first, so as to destroy the context of the two adjacent dorsal consonants. This would be a bleeding order. The correct form, however, is [slɑŋskə]. That is, we need to apply the rules in a counterbleeding order: first (21), then (20).

(22)  bak-kə  sl着重 n-kə  snɔr-kə

Rule (21)  s  s  (n.a.)
Rule (20)  (n.a.)  $\emptyset$  $\emptyset$

Q126 Mwera has three rules, given as (1), (2) and (3) below. Two noun
stems in Mwera are [gomo] ‘lip’ and [kuja] ‘cape bean’. The plural is
formed by prefixing a nasal consonant, whose underlying form is [n]. The plural surface forms are [ŋomo] and [ŋuja]. The following
three rules derive the surface forms (Kenstowicz and Kisseberth

| Rule (2) | [−son] |
| Rule (3) | [−cont] \rightarrow [−\text{nas}] |

+ voice
Q127 Dutch has a number of rules affecting the feature [±voice] in obstruents. In order to derive the surface forms in the second column from the underlying forms in the first, four rules are required: \textsc{degemination}, \textsc{final devoicing}, \textsc{progressive devoicing} and \textsc{regressive voicing}. The first two can be formalized as follows:

\begin{itemize}
\item \textsc{degemination} \quad \begin{array}{c}
\begin{array}{c}
[+\text{cons}] \\
[\#]
\end{array}\quad \rightarrow \quad \begin{array}{c}
\begin{array}{c}
Ø
\end{array}
\end{array}
\end{array}
\quad \text{Condition: } 1 = 3
\end{itemize}

\begin{itemize}
\item \textsc{final devoicing} \quad [−\text{son}] \rightarrow [−\text{voice}] \quad / \quad C_0 \#
\end{itemize}

\begin{tabular}{lll}
\textbf{Underlying} & \textbf{Derived} & \\
\hline
a & la:t#bluj\textcircled{r} & la:bluj\textcircled{r} \quad \text{‘late developer’} \\
b & bad#bruk & badbruk \quad \text{‘bathing trunks’} \\
c & kaud#vy:r & kautfy:r \quad \text{‘gangrene’} \\
d & vand#te:\textcircled{\textbf{c}}\textcircled{\textbf{e}}l & vante:\textcircled{\textbf{c}}\textcircled{\textbf{e}}l \quad \text{‘wall tile’} \\
e & le:z#fa\textcircled{\textit{ut}} & le:sf\textcircled{\textit{ut}} \quad \text{‘reading error’} \\
f & lup#z\textcircled{\textit{e}}\textcircled{\textit{y}}\textcircled{\textit{v}}\textcircled{\textit{r}} & lup\textcircled{\textit{s}}\textcircled{\textit{e}}\textcircled{\textit{y}}\textcircled{\textit{v}}\textcircled{\textit{r}} \quad \text{‘very pure’} \\
g & bo:t#t\textcircled{\textit{x}}\textcircled{\textit{t}} & bo:t\textcircled{\textit{x}}\textcircled{\textit{t}} \quad \text{‘boat trip’} \\
h & ka:z#za:k & ka:sa:k \quad \text{‘cheese shop’} \\
i & kop#bal & k\textcircled{\textit{b}}\textcircled{\textit{a}}l \quad \text{‘header’ (football)} \\
j & \text{\textcircled{\textit{xa}}\textcircled{\textit{ud}}#del\textcircled{\textit{v}}\textcircled{\textit{r}} & \text{\textcircled{\textit{xa}}\textcircled{\textit{ud}}\textcircled{\textit{el}}\textcircled{\textit{v}}\textcircled{\textit{r}} \quad \text{‘prospector’} \\
k & \text{\textcircled{\textit{xa}}\textcircled{\textit{ud}}#ko:ts & \text{\textcircled{\textit{xa}}\textcircled{\textit{ut}}ko:ts \quad \text{‘gold fever’} \\
l & krab#s\textcircled{\textit{xa}}:v & kraps\textcircled{\textit{xa}}:f \quad \text{‘scraper’} \\
m & le:z#bril & le:z\textcircled{\textit{br}}\textcircled{\textit{il}} \quad \text{‘reading glasses’} \\
n & heis#balk & heizbalk \quad \text{‘hoisting beam’} \\
\end{tabular}

1\quad \text{Determine what the other two rules should do, and how the four rules should be ordered.}
2\quad \text{Give formal notations of \textsc{progressive devoicing} and \textsc{regressive voicing}.}
3\quad \text{Give the derivations of items (d), (f), (h), (i) and (m).}
Q128 The river *Linge* [lɪŋə] flows not far from Kaatsheuvel. What would you expect the diminutive form to be in that dialect?

Q129 In (10), the order in which **flapping** and **pre-fortis clipping** are applied to *rider* and *writer* in mainstream Canadian English results in different surface forms for these two words. What type of rule order is this?

### 13.4 TRANSPARENCY AND OPACITY: RULES AND CONSTRAINTS

Which of the rule orderings discussed in the previous section are the more natural, or more expected? The answer depends to some extent on the perspective that is taken. Kiparsky (1968) has argued that historical change could in part be explained by assuming that over time rules tend to reorder so as to maximize their application. If rule A feeds rule B, the number of forms to which rule B can apply is increased, because rule A supplies additional forms to rule B, as was the case for *preglottalization*, discussed earlier. As a result, rule B can be maximally applied. Obviously, in a counterfeeding order, rule B will not be supplied with additional forms to which it can apply, as illustrated for French *final consonant deletion* in 13.3.2. Similarly, if rule A removes forms to which rule B can potentially apply (A – B: a bleeding order), the opposite order, counterbleeding, allows rule B to apply maximally. Thus, on the basis of the principle of **maximal rule application**, feeding and counterbleeding are the natural rule orders. Whereas there was general consensus that feeding was more natural than counterfeeding, there was disagreement as to whether bleeding or counterbleeding should be considered the more natural order.

But what determines what is ‘natural’ or ‘unmarked’? From the perspective of rule maximization, feeding and counterbleeding are natural rule orders, but an alternative perspective was that rules are more natural when their application is transparent on the surface. From the point of view of **maximal surface transparency**, bleeding and feeding are the natural orders. To see this, two more cases of counterfeeding and counterbleeding rule orders are presented.

In Gran Canarian Spanish (Oftedal 1985), the voiced stops /b d ɡ/ are spirantized to [β δ ɣ] in intervocalic position, as shown in (23a). Additionally, *Voicing* causes intervocalic /p t k/ to be realized as [b d ɡ], as illustrated in (23b).

(23a) Underlying | *Spirantization* | Sense
---|---|---
/roba/ | [roβa] | ‘he/she steals’
/nada/ | [naða] | ‘nothing’
/la gana/ | [la ɣana] | ‘the appetite’
A counterfeeding application of Spirantization followed by Voicing straightforwardly generates the correct surface forms. However, from a surface perspective, the voiced plosive in [la gama] does not make good sense, given the existence of Spirantization. By this rule, intervocalic /g/ surfaces as [ɣ], but [g] survives in surface forms like [la gama].

Or, again, Slovak (Rubach 2000) has a diminutive suffix /æ/ which triggers Palatalization of a preceding stem-final coronal, as in páňa /pan-æ/ [paňa] ‘master (DIM)’. A further process of æ-Backing backs underlying /æ/ to [a] after non-labials, by which [paňa] becomes surface [pana]. So, when the diminutive suffix /æ/ is preceded by a labial consonant, as in holúbľa /holub-æ/ [holubá] ‘pigeon (DIM)’, neither rule applies. This is an example of a counterbleeding rule order: Palatalization applies before æ-Backing to prevent /pan+æ/ from surfacing with incorrect *[pana]. From the perspective of maximal rule application, this counterbleeding rule order is natural, since it makes Palatalization maximally applicable, but from the perspective of maximal surface transparency it is not, since it makes Palatalization a non-transparent rule. This is because the surface form [pana] has a palatalized consonant in the absence of a following front vowel that triggers the palatalization. In other words, something happens, but it is unclear (opaque, not surface transparent) why it does. For a counterfeeding rule order, things are the other way around: something does not happen ([la gama] not turning into [la ýama]), but it is not clear (not surface transparent) why it does not. That is, both counterfeeding and counterbleeding lead to surface opacity. Counterbleeding leads to forms with rule outputs that have no context and is therefore also referred to as over-application. Counterfeeding leads to forms with contexts that spare a target and is also referred to as under-application. These two opaque rule interactions are problematic for classic OT, whose constraints apply to surface forms.

13.5 OPACITY AND OT: TWO SOLUTIONS

13.5.1 Counterfeeding opacity is problematic for OT

OT has no problems in accounting for feeding (transparent) rule orders, but cannot handle counterfeeding (opaque) rule orders. This is illustrated for Gran Canarian Spanish in tableau (25), which contains two faithfulness constraints and two markedness constraints, as listed in (24).

(24) Ident(voice): A consonant in the output has the same [voice] specification as in the input.
Identifier(continuant): A consonant in the output has the same [continuant] specification as in the input.
Voicing: No voiceless intervocalic plosive.
Spirantization: No voiced intervocalic plosive.
Tableau (25a) shows that, in order for Spirantization to take place, the constraint Spirantization must dominate Ident(cont). If it does not, the first output candidate will be selected as optimal. Similarly, as shown in (25b), to produce voiced intervocalic plosives, Voicing must dominate Ident(voice). This ranking, however, incorrectly predicts that /la kama/ is optimally realized as [la ɣama] rather than as actual [la gama].

Q130 Show that the ranking of Voicing » Ident(voice) below Spirantization » Ident(cont) does not produce the correct results either. What is the incorrect result of this ranking for the intervocalic plosives?

Tableaux (25a) and (25b) show that OT cannot account for a counterfeeding rule interaction. As it happens, counterfeeding situations involving spirantization and voicing are quite common. The Gran Canarian Spanish case can also be observed in Sardinian (Bolognesi 1998) and Corsican (Gurevich 2004). Kaye (1975) proposed a functional motivation for counterfeeding opacity: keeping intact lexical contrasts and thus avoiding neutralization. Indeed, if spirantization and voicing were to stand in a feeding relationship, the surface forms for minimal pairs like /nada/ [naða] and /nata/ [nada] ‘cream’, or /roba/ [roβa] and /ropa/ [roba] ‘clothes’, would be realized identically, as [naða] and [roβa] respectively, neutralizing the underlying contrast. In fact, Gurevich (2004) observes in her study of 230 similar processes in a corpus of 153 languages that in the majority of cases (92%) they are non-neutralizing. Incidentally, this contrast-preserving tendency is not only problematic for OT, but also for a rule-based theory. Hale and Reiss (2008: 14) state: ‘Opaque rules are not surface true, rules that are not surface true are harder to learn, failure to learn aspects of the ambient language constitutes a diachronic change, therefore, languages are more likely to lose a given instance of opacity than to gain one.’ Now, if the unmarked rule order is a feeding rule order (Kiparsky 1982b), it is not easy to understand why in the majority of cases a counterfeeding ordering between Spirantization and Voicing is attested.
Counterbleeding opacity is problematic for OT

A counterbleeding rule order of Palatalization and æ-Backing leads to a palatalized consonant in Slovak [paɲa] for underlying /pan+æ/ ‘master (dim)’. However, a palatalized consonant here appears without its context /æ/ and as such is not transparent. Tableau (27), where the faithfulness and markedness constraints in (26) are used, shows that counterbleeding opacity is as problematic for OT as counterfeeding opacity. The third output candidate is incorrectly selected as optimal, instead of the actual Slovak form [paɲa].

(26) Ident(back): A vowel in the output has the same value for the feature [back] as in the input.
Ident(ant): A consonant in the output has the same value for the feature [anterior] as in the input.
Palatalization: No non-palatalized coronal consonants before a front vowel.
æ-Backing: No [æ] after non-labials.

(27) /pan+æ/ æ-Backing Ident(back) Palatalization Ident(ant)
panae *!
panæa *!
æ paña *
pana * *!

Q131 Show that a ranking æ-Backing » Ident(back) below Palatalization » Ident(ant) is not helpful in obtaining a counterbleeding relationship between Palatalization and æ-Backing.

The obvious reason why OT cannot describe counterfeeding and counterbleeding opacity is the use of output markedness constraints. A constraint such as Spirantization (no voiced intervocalic plosive) cannot differentiate between [la gana] as an output of /la gana/ and [la gama] as a possible output of /la kama/, neither of which satisfy Spirantization. A rule-based approach can exploit the underlying difference between /g/ in /gana/ and /k/ in /kama/ by applying Spirantization before Voicing, creating [la gama] from /la kama/ after Spirantization has applied. In OT, there is no straightforward way to make use of that underlying difference, because its output constraints cannot refer to inputs. In the next section, we will briefly discuss the merits of two solutions that have been proposed to allow OT to handle counterfeeding and counterbleeding opacity.

Stratal OT and bringing some derivation back in the OT model

In chapter 8, we have seen that there are good reasons to assume that, in addition to the underlying and the surface representation, there is an intermediate level of representation: the lexical representation. That idea has also been implemented in OT
and is generally known as **Stratal OT**. The underlying form passes through a lexical constraint hierarchy and yields an optimal intermediate output form. This form is then input to a second constraint hierarchy, potentially different from the first, and produces the optimal surface form (Booij 1997; Rubach 2000; Kiparsky 2015; Bermúdez-Otero forthcoming). For counterbleeding, opaque Slovak **palatalization**, Rubach (2000) has proposed the following analysis. In the first constraint hierarchy, the order of the constraints æ-**backing** and IDENT(back) is the opposite to that in tableau (27). This is illustrated in (28), where the optimal output candidate now becomes the second output candidate, [paña], instead of [pana].

This optimal intermediate output form [paña] is then input to a second constraint hierarchy, where the order of the constraints æ-**backing** and IDENT(back) is inverted, as illustrated in (29).

The result of this two-level approach is that correct [paña] is selected as the optimal output form for underlying /pan+æ/.

### Q132 Provide a Stratal OT analysis of the counterfeeding case in Gran Canarian Spanish, such that **spirantization** is satisfied at the first level, but **voicing** is not.

Although Stratal OT can describe the problematic non-surface-transparent interactions, it is not always evident that the different levels are independently motivated. **Spirantization** and **voicing** would at first sight seem to belong to the postlexical domain. **Spirantization** would be considered a postlexical process in the model outlined in chapter 8, because it produces a novel segment, [ɣ]. If the first level in Q132 is a postlexical level, a second postlexical level is required for **voicing** to be satisfied by non-spirantized plosives.

Moving on the second solution, we note again that what needs to be expressed is that it is all right for Gran Canarian Spanish /la kama/ to change into [la gama], but
that it should not further change into [la ɾama]. Similarly, Slovak /pan-æ/ should
first change to [paɾæ] and then to [paɾa], but not directly from /pan-æ/ to [paɾa].
To achieve this, it is necessary to keep track of the derivational history. Standard OT
cannot do that, as markedness constraints are output constraints. As such, they can
refer neither to the input nor to the way input and output are connected. The second
solution, then, is to make sure that changes in the phonological representation take
place one by one, rather than at one go. A change in the phonological representa-
tion of a word implies that a faithfulness constraint is violated, but a markedness
constraint respected. Going from input to output, successive changes produce a
series of intermediate forms that step-wise give up faithfulness in order to improve
on markedness. In (30), we illustrate this for Slovak /pan+æ/. The input is given
in the first column, and the second column lists the chain of intermediate forms,
whereby the faithfulness constraint which an intermediate form violates is listed
below the intermediate form in question.

(30)

<table>
<thead>
<tr>
<th>Input</th>
<th>Intermediate forms</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 /pan-æ/</td>
<td>&lt;panæ&gt;</td>
<td>[paɾæ]</td>
</tr>
<tr>
<td></td>
<td>No Faithfulness Violations</td>
<td></td>
</tr>
<tr>
<td>2 /pan-æ/</td>
<td>&lt;panæ&gt;, &lt;paɾæ&gt;</td>
<td>[paɾæ]</td>
</tr>
<tr>
<td></td>
<td>IDENT(ANT)</td>
<td></td>
</tr>
<tr>
<td>3 /pan-æ/</td>
<td>&lt;panæ&gt;, &lt;paɾa&gt;</td>
<td>[paɾa]</td>
</tr>
<tr>
<td></td>
<td>IDENT(BACK)</td>
<td></td>
</tr>
<tr>
<td>4 /pan-æ/</td>
<td>&lt;panæ&gt;, &lt;paɾæ&gt;,</td>
<td>[paɾa]</td>
</tr>
<tr>
<td></td>
<td>&lt;paɾa&gt;</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IDENT(ANT) IDENT(ANT)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IDENT(BACK)</td>
<td></td>
</tr>
</tbody>
</table>

Chain 1 just contains the fully faithful form. Chain 2 contains an interme-
diate form that violates IDENT(ANT), while chain 3 contains one that violates
IDENT(BACK). Chain 4 contains two intermediate forms, <paɾæ>, which violates
IDENT(ANT), and <paɾa>, which violates both IDENT(ANT) and IDENT(BACK).
How can we now implement the idea that /pan+æ/ should first change to [paɾæ] and
then to [paɾa], but not directly from /pan+æ/ to [paɾa]? This can be done by impos-
ing an order on the violations of the faithfulness constraints. For this, a new type of
constraint is introduced, PREC(EDENCE), which is given in (31), where A and B are
faithfulness constraints.

(31) PREC(A,B): Any form violating B is preceded by a form violating A and is not followed by a form violating A.

There are thus two ways of violating (31): by having a B-violator without a
preceding A-violator in the chain and by a B-violator which is followed by an A-violator in the chain. To check whether an output candidate violates
PREC(IDENT(ANT),IDENT(BACK)), say, we scan the list of intermediate forms to
see whether a form that violates IDENT(BACK) is preceded by a form that violates
IDENT(ANT). If this is not the case, a violation of PREC(IDENT(ANT),IDENT(BACK))
is incurred. Next, if a form that violates IDENT(BACK) is followed by one that
violates \textsc{Ident(ant)}, another violation is incurred (McCarthy 2007). For Slovak, we illustrate this in (32), which lists the output candidates in the first column, as usual, but now together with the input form and the chain of intermediate forms. Constraint \textsc{Prec(Ident(ant),Ident(back))} is irrelevant for candidate 1, which violates neither of the two faithfulness constraints, while candidate 2 satisfies it, because the form with the violation of \textsc{Ident(ant)}, \texttt{<panæ>}, is not preceded by a form that violates \textsc{Ident(back)}. However, \textsc{Prec(Ident(ant),Ident(back))} is violated by candidate 3, which has a form violating \textsc{Ident(back)} in its chain of intermediate forms which is not preceded by one showing a violation of \textsc{Ident(ant)}. Candidate 4 violates both \textsc{Ident(ant)} and \textsc{Ident(back)}. A scan of the chain of intermediate forms shows that the violations occurred in the right order, and there is thus no violation of \textsc{Prec(Ident(ant),Ident(back))} by candidate 4.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Input & Intermediate forms & Output \\
\hline
1 /\texttt{la gana}/ & \texttt{<la gana>} & \texttt{[la gana]} \\
& No Faithfulness Violations & \\
\hline
2 /\texttt{la gana}/ & \texttt{<la gana>, <la kana>} & \texttt{[la kana]} \\
& \textsc{Ident(voice)} & \\
\hline
3 /\texttt{la gana}/ & \texttt{<la gana>, <la ãana>} & \texttt{[la ãana]} \\
& \textsc{Ident(cont)} & \\
\hline
4 /\texttt{la kama}/ & \texttt{<la kama>, <la gama>} & \texttt{[la gama]} \\
& \textsc{Ident(voice)} & \\
\hline
5 /\texttt{la kama}/ & \texttt{<la kama>} & \texttt{[la kama]} \\
\hline
6 /\texttt{la kama}/ & \texttt{<la kama>, <la gama>, <la ãama>} & \texttt{[la ãama]} \\
& \textsc{Ident(voice)} \quad \textsc{Ident(voice)} \quad \textsc{Ident(cont)} & \\
\hline
\end{tabular}
\caption{Gran Canarian Spanish can be brought in line with \textsc{Prec(Ident(cont), Ident(voice))}. In (33), the chains are listed in the same format as in (30).}
\end{table}
In (34a), the treatment of /la gana/ is shown. As it happens, the Precedence constraint has no crucial role in accounting for the exclusion of candidate 2. While the chain of intermediate forms actually contains a form violating Ident(voice) without a preceding form with a fricative, the violation of Prec(Ident(cont),Ident(voice)) is irrelevant. Thus, OT with chains is only relevant to ‘harmonic’ forms, i.e. candidates that are not already eliminated by constraints that outrank the Precedence constraint, like candidate 2 in (34a). Candidate 1 violates neither faithfulness constraint, so that again Prec(Ident(cont),Ident(voice)) has no role to play, but founders on Spirantization. As a result, the uncontested winner is candidate 3.

We thus move on to /la kama/. Because candidate 5 is eliminated by voicing, the competition for /la kama/ is between candidates 4 and 6, as shown in tableau (34b). Candidate 6 violates Prec(Ident(cont),Ident(voice)) twice. First, the violation of Ident(voice) is not preceded by a violation of Ident(cont) and, second, the violation of Ident(voice) is followed by a violation of Ident(cont). The violation by winning candidate 4 of Spirantization is non-fatal, since its fate as the winner is ensured by obeying the Precedence constraint.

In this section, we have briefly illustrated two ways in which OT allows us to describe counterfeeding and counterbleeding opacity. The simpler solution was that of Stratal OT, which assumes different phonological grammars for different stages.
in the process of the incorporation of phonological forms into higher levels of the morphosyntactic structure. A disadvantage was that in the case of Gran Canarian Spanish the two levels could not be independently motivated. The second solution involved the addition of a mechanism that keeps track of the history of changes to an underlying representation. An evaluation of OT with chains will need to recognize that this theory imports some form of ‘extrinsic’ order into a theory with freely rankable constraints. If the set of constraints is truly universal, as we mentioned in chapter 4, this implies that all the Precedence constraints required to describe all possible opaque interactions will have to be part of the constraint set. Together with the lack of independent evidence for the mechanics of the solution, this implication may leave one with an expectation of new and better insights.

13.6 CONCLUSION

In this chapter we have discussed and exemplified four possible types of rule interaction: feeding, counterfeeding, bleeding and counterbleeding. We established that counterfeeding and counterbleeding lead to opaque phonological surface forms. These contradict some phonological rule of the language, either by showing the effect of the rule where there is no longer any context for it (counterbleeding), or by not showing the effect where the context is present (counterfeeding).

We have also seen that OT cannot account for opaque forms if markedness constraints are there to say something about surface forms. In response to the inherently ‘transparent’ nature of OT constraints we briefly reviewed two solutions: Stratal OT and OT with chains of intermediate forms connecting inputs to outputs. The basic observation made in chapter 4 was that one of the goals of phonological theory is to account for the fact that morphemes may be pronounced differently depending upon their environment. When we compared rule-based and constraint-based descriptions, we argued that OT is to be preferred, given that it does not run into the duplication problem – the necessity of stating the same generalization both in a phonological rule and in a constraint, a problem inherent in rule-based descriptions. Also, OT was argued to be able to express directly how any new input forms from any donor language are brought in line with the structural demands of the recipient language. Systematic modifications in pronunciation thus seemed to be taken care of by OT. In this chapter, we have seen that OT can deal with opacity, a common fact of life in the phonologies of languages, but only by introducing additional machinery that does not always seem well motivated.

This is the point where we should step back and briefly evaluate the phonological enterprise. Since SPE, theoretical phonology has had a two-fold research agenda. One of these focused on representations and has led to the introduction of moras, syllables, feet and more, as we have shown in the preceding chapters. The second has focused on levels of representation and how these are connected. While the first focus has led to exciting insights, the second focus on the phonological research agenda leaves us wondering what went wrong. Without much ado, a speaker of Gran Canarian Spanish will take a sound sequence [la gama] to refer to a
representation /la kama/, just as he will take a sound sequence [la ɣama] to refer to a representation /la gama/. But none of the theories that we have discussed would appear to convincingly explain how underlying /la kama/ and /pan+æ/ get to be [la gama] and [paɲa] in actual pronunciation. Perhaps it is relevant to note that, whether rule-based or constraint-based, all input-output models appear to have a speaker orientation in common, a perspective that might be too narrow. There have been proposals that take the opposite perspective (Smolensky 1996; Boersma 1998; Kager 1999). Progress in this area could come from research into the way language production and perception are dealt with in real life, in the language user’s brain. Such research is now being conducted both by psycholinguists on the basis of various behavioural tasks and by neurocognitive scientists on the basis of the time course and localization of brain activity associated with producing and perceiving speech. New perspectives on the old problem of connecting inputs and outputs may well therefore be waiting around the corner.
In this last chapter, we will present a hierarchical configuration of the distinctive features introduced in chapters 5 and 6. The main advantages of this autosegmentalized representation are, first, that features can spread to neighbouring segments, thereby representing the partial identity of adjacent segments in the most parsimonious way, and, second, that specific groups of features can spread, thereby making the claim that not all combinations of features can be the focus of an assimilation. As for the first point, assimilations show a non-arbitrary relationship between the context and the target. Time and again, rules appear to transfer a specific feature or group of features from one segment to a neighbouring segment. However, the representation we introduced in chapters 5 and 6 is incapable of expressing this. Consider once more the Dutch rule of regressive voicing, which voices obstruents before [b d]. It is a natural rule in the sense that voiced segments cause preceding segments to be voiced, as opposed to voiceless or labial, say. An arrangement of segments consisting of self-contained lists of features will not provide a natural way of expressing the transfer of features from one segment to the next. Neither can we make a distinction between the transfer of an expected feature, [+voice], in regressive voicing, and an unexpected feature. An elegant solution to this predicament should allow segments to share the same feature, much in the way successive tone-bearing units can be associated with the same tone. As for the second point, in order to transfer the right groups of features from one segment to the next, the representation should allow only those feature groups to be shared that actually go together in assimilation rules. Many languages have processes that transfer the features specifying a consonant’s place of articulation to a preceding nasal consonant, for instance. If the features are unstructured lists, it is impossible to distinguish between groups of features that are never shared and groups that routinely are. That is, we would not want an unnatural feature group like [labial, ±sonorant] to be as easily transferrable as a natural one like [±voice, ±spread glottis, ±constricted glottis].

In feature geometry (Clements 1985, 1993; Sagey 1986; McCarthy 1988), the problem of the grouping of features is solved by representing segments as trees, in which the nodes represent features and feature groups. The sharing problem is solved by assuming that a single node may be part of more than one tree. To show the merits of this proposal, two main topics will be dealt with. One concerns the opportunity that is afforded by the underspecification of features – and, hence, the absence of particular nodes in the representations of segments – to characterize long-distance assimilations. Particularly relevant in this context are word-based
restrictions on the distribution of specific features, as is seen in vowel harmony. The solution to characterizing these distributional restrictions is reminiscent of the description of the lexical tone patterns of languages like Etung and Mende, discussed in chapter 10. The second topic concerns the representation of the three types of complex segments recognized in chapter 2. Finally, we will return to the question of the representation of palatal-alveolar and palatal segments, like [tʃ] and [j], and argue that they, too, are complex segments.

14.2 NATURAL FEATURE CLASSES

As observed above, assimilations of place involve the transfer of a collection of features. Many languages have rules assimilating nasal consonants to the place of articulation of the following consonant. For instance, Hindi, which has bilabial, dental, retroflex, palato-alveolar and velar oral stop consonants, also has the corresponding nasals [m n ŋ n], whose place of articulation is determined by the following consonant in the word. Word- Internally, sequences like *[nk], *[ŋt], *[ŋc] or *[mt] are therefore ill-formed. As a result, a prefix like [sam] ‘together’ shows alternations of the type illustrated in (1) (Ohala 1983). Clearly, the place features must be accessible as a group in order to express this phonological process, the transfer of all the place features, including dependent ones like [+anterior], as shown by the words for ‘equilibrium’ and ‘movement’.

(1) aːkaːr ‘shape’ samaːkaːr ‘homophonous’
    kiːrtan ‘devotional singing’ saŋkiːrtan ‘collective devotional singing’
    joːl ‘measure’ səŋjoːl ‘equilibrium’
    calan ‘conduct’ səŋcalan ‘movement’
    naːd ‘sound’ sannaːd ‘consonance’

There are also processes that transfer a subgroup of the place features. English CORONAL ASSIMILATION assimilates [t d n l] in place of articulation to the following coronal consonant. That is, before [θ ð] they are dental, before [t d n l] they remain alveolar, and before [a], which in English has a postalveolar place of articulation, they are postalveolar. In (2), we repeat the feature values for the segments concerned. The data in (3) illustrate how the values of the features [anterior] and [distributed] are together passed on to the preceding [−cont, +cor] segment (Clements 1985).

(2) | COR | Ant | Distr |
--- | --- | --- |
θ ð n | ✓ | + | + |
t d n l | ✓ | + | – |
ʃ l | ✓ | – | – |

(3) in Rome
    ɪŋ ɹoʊm
    ɔːl̪ðεə ‘all there’
    in Thirsk
    ɪŋ ˈθɜːsk
    try
    ɡɛt ɹɛd
    get red
    wɪdθ
    width
Of course, single features are also frequently transferred from one segment to the next. Consider the Old English data in chapter 7 (Q61(2)), which showed that the [-back] [ç] appeared after [-back] vowels, and [+back] [x] after [+back] vowels. This process is understandable if we assume that the vowel transfers the value for the feature [back] (but no other features) to the following dorsal fricative. Just as phonological processes are seen as evidence for the existence of natural segment classes (any group of segments referred to by a process), so a natural feature class can be defined as a group of features that is manipulated (i.e. transferred, deleted or inserted) by some phonological process.

14.3 BUILDING A TREE

Under a theory in which the representation of a segment is an unstructured list of features, the representation of [t] would be as in (4a). What rules like Hindi nasal place assimilation suggest is that the representation must be something like (4b). Further, on the basis of processes like that illustrated in (3) we would seem to need a representation like that in (4c). What this means, of course, is that the segment is not a single constituent, but has internal structure. Such constituency is best represented by means of a tree diagram.

14.3.1 The place node

The place node dominates articulator nodes corresponding to the univalent features [labial], [coronal], [dorsal] and [radical], as seen in (5). Each of these dominates subconstituents corresponding to their relevant features. Thus, [labial] dominates [±round], [coronal] dominates [±anterior] and [±distributed], and [dorsal] dominates [±back], [±high] and [±low]. With the help of this representation, processes can refer to the constituent [labial, distr, round], but not to
the nonconstituent [high, ant], for instance. We will assume that the position of \([\pm \text{tense}]\) will be under \([\text{RAD}]\), assuming it is equivalent to \([\pm \text{ATR}]\).

\[
\begin{align*}
\text{LABIAL} & \quad \text{COR} & \quad \text{DORSAL} & \quad \text{RAD} \\
\text{round} & \quad \text{ant} & \quad \text{distr} & \quad \text{back} & \quad \text{high} & \quad \text{low} & \quad \text{tense}
\end{align*}
\]

14.3.2 The laryngeal node

What other natural feature classes are there besides place features? Lass (1976: 145ff.) has argued that the laryngeal segments \([\text{h} \ ?]\) have no supralaryngeal specification at all. He based his case on the frequently observed phenomenon that obstruents sometimes appear to lose all their supralaryngeal information. Consider the London English data in (6), which are representative of this kind of process. In certain contexts, preglottalized plosives are pronounced as \([\text{ʔ}]\). (The variation between preglottalized plosives and glottal stops is stylistic; i.e. both pronunciations may be heard from the same speaker.)

\[
\begin{align*}
\text{Stage I} & \quad \text{Stage II} \\
\text{mismo} & \quad \text{miismo or miismo} & \quad \text{‘same’} \\
\text{fosforo} & \quad \text{fosforo} & \quad \text{‘match’}
\end{align*}
\]

In New World Spanish, \([\text{s}]\) has been lost in coda position. That is, the plural of \([\text{klase}]\) ‘class’, which in Peninsular Spanish is \([\text{klases}]\), is \([\text{klasε}]\). (The \([\varepsilon]\) arises from a separate process which is not relevant here.) However, the \([\text{s}]\) has not left without a trace: an \([\text{h}]\)-like offglide can be heard after the plural in careful speech. Similarly, voiceless sonorants have arisen word-internally after the loss of the oral articulation of \([\text{s}]\), as shown in (7).

Both processes can be described if laryngeal and supralaryngeal features are separate constituents. Under that assumption, both the English and the Spanish data would receive the same description: deletion of the supralaryngeal constituent. In the case of English, this would leave \([+\text{constr}, -\text{voice}, -\text{spread}]\) behind, i.e. \([?]\), and in the case of Spanish it would leave \([-\text{constr}, -\text{voice}, -\text{spread}]\) behind, i.e. a voiceless interval at that point in the word. Such orphaned laryngeal features thus generally end up as \([?]\) and \([\text{h}]\), respectively. To express Lass’s proposal in our feature tree, we need to assume that segments have a LARYNGEAL and a SUPRALARYNGEAL
To ensure that segments with only one of these constituents also have specifications for the major-class features ([±consonantal] and [±sonorant]), these features must be present in both constituents. According to McCarthy (1988), they make up the root node. This decision can safely be made, since no processes have been reported that transfer [±consonantal] or [±sonorant] from one segment to the next. This means that our tree now looks like (8). We will treat the feature [±approximant] as a manner feature.

(8)

14.3.3 The supralaryngeal node

In the preceding sections, we have been able to establish a constituent structure for the major-class features, the laryngeal features and the place features. What remains to be determined is the location of the manner features [±nasal], [±continuant] and [±lateral], in addition to [±strident] and [±approximant]. Do they form a third constituent, or are they located under either the laryngeal or the supralar/place node? We argue that [continuant] is not located in the laryngeal constituent in (8) on the basis of an assimilation process changing [h] to [ç] in Frisian (Tiersma 1985). Recall from chapter 5 that [h?] are not specified for [±cont], and that this was the explanation for the fact that they do not trigger nasalization. This rule nasalizes a vowel before [n], provided a [±cont] consonant follows. In (9a) the effect of the rule is illustrated, while (9b) shows that [h] does not trigger it.

(9)  

Knowing that [±continuant] is not in the left-hand branch of (8) is not enough to decide whether it is dominated by the supralar/place node or directly by the root node. The southern variety of Frisian has a rule of diphthong shift, which – among other vocalic changes – turns [iə] into [jε]. When this [j] appears after [h], it assimilates [h] to [ç]. That is, the place features of [j] are added to the representation of [h]. The question now is whether [±continuant] is transferred along with the place features. If it is dominated by the supralar/place node, [±cont] of [j] will be transferred to [h], along with its place features. If it is included as a third branch under the root node, [h] would only receive a place specification, but would
continue to lack the specification [+cont]. That is, transfer of [+cont] predicts that after the assimilation, [ç] will trigger nasalization, while no transfer predicts that [ç] will not trigger nasalization. The data in (10) show that the new [ç] triggers nasalization and must therefore be [+cont]. This segment is now [–son, –cons, +spread, –voice] (which is what [h] contributes), as well as [+cont, –lat, –nas, cor, –ant] (which is what [j] contributes).

(10) Underlying | Northern   | Southern       |
--- | --- | --- |
inhɪəkə | inhɪəkə | ɪ̂çjεkə | ‘to hook in’
oən-hɪərə | oən-hɪərə | oə̃çjεrə | ‘to listen to’

The tree in (8) predicts that segments other than [h ɦʔ] can lose or acquire manner features independently of the laryngeal features. This can be shown for the lateral consonants of Klamath. In addition to a voiced [l], Klamath has a glottalized [lˀ] and voiceless [l̥]. When appearing after [n], a lateral consonant transfers its lateral articulation to the preceding segment, but without its laryngeal features, which remain behind as [ʔ] and [h], as shown in (11). This again shows that there is a constituent in the feature tree that contains manner features but not the laryngeal features. Since the new segment is not [+nas], this Klamath process suggests that the supralar/place node of (8) contains [±lateral] (because it is transferred from the second to the first segment) as well as [±nasal] (because it is deleted from the first segment), on the correct assumption that these features are mutually exclusive in the language.

(11) [nlˀ] → [lʔ]
     [nl̥] → [lh]

Q133

1. Which constituent is transferred from a consonant to a preceding [t d n l] by English coronal assimilation?
2. Which constituent is transferred from a consonant to a preceding nasal by Hindi nasal assimilation?

While it is therefore clear that the manner features are inside the supralar/place constituent, it is also clear that they cannot be included under the place node. In Hindi, as in many other languages, it is just the consonant’s place node that is transferred to a preceding nasal, not its [–nas] specification. From the node labelled supralar/place in (8), therefore, we must split off the place node. Since there have been no reports of processes that transfer [±nasal], [±continuant], [±lateral], [±strident], [±approximant], en bloc, from one segment to the next, these
features are not grouped in a manner constituent, but form separate terminal nodes
dominated by supralar, as in (12).

(12)

Single manner features can be shown to transfer independently to neighbouring
segments. Processes that transfer [+nas] were discussed for Frisian in this
section and English in section 8.6.2. Transfer of [±continuant] is much rarer. In
American English, forms like [gɪbm] for given, [hiːd̪n̩] for heathen, [bɪdnəs] for
business and [wʌdn̩t] for wasn’t occur (Bailey 1985: 63). Transfer of [±lateral]
occur in Klamath, as seen above. No case has been made for the transfer of
[±strident] or [±approximant]. There have been a number of proposals that modify the structure of (12); however, (12) may be seen as a ‘consensus’ tree
(Broe 1992).

14.4 SPREADING AND DELINKING

In this section, we will deal with the way features or subconstituents of the seg-
ment tree spread or delink. In our representation, segments are arranged from left
to right, which is the usual metaphor for the order from ‘early’ to ‘late’. Along this
time axis, the segment trees are arranged much in the way the records in an old-
fashioned jukebox are arranged: a row of parallel disks, where a disk represents a
segment. Each disk defines a plane in which some instantiation of the feature tree is
to be found. If you were to take out a disk to look at its right or left face, you would
see some version of (12). This is known as an end view.

How can a node of some segment be dominated by the appropriate node of an
adjacent segment? Imagine two adjacent segments S1 and S2. Now mentally draw
a line from the [±nasal] node of S1 to the [±nasal] node of S2, and a line from the
supralar node of S1 to the supralar node of S2. The lines you have just drawn
are, of course, tiers. All corresponding nodes form tiers in this way, so that we can
talk about the ‘nasal tier’, the ‘supralaryngeal tier’, the ‘root tier’, etc. Evidently, tiers
are adjacent if no tier intervenes between them. This is the case with the place tier
and the supralaryngeal tier, for instance: no node intervenes between them. Two
adjacent tiers define a plane.
Q134 Draw two feature trees in pencil, next to each other. Draw all tiers. Now erase all pencil lines that are ‘behind’ the planes nearest to you. As you will realize, your picture shows three dimensions.

We can now draw, in some plane, an association line from any node to the immediately dominating node in an adjacent segment. That is, from the [labial] node of \( S_1 \) we can draw an association line to the place node of \( S_2 \), and from the place node of \( S_1 \) to the supralaryngeal node of \( S_2 \), and so on. This spreading of nodes in the feature tree is entirely comparable with the spreading of tones, or with the association of one vowel to two moras.

Q135 Can you draw an association line from the place node of \( S_1 \) to the root node of \( S_2 \)? If not, why not?

In Hindi place assimilation, the place node of the right-hand segment associates with the supralaryngeal node of the left-hand segment. This is shown in (13) by the interrupted association line. Notice that there is something wrong now: what was an \([n]\) is now specified for two sets of place features. What we need to do is delink the original specification. As usual, this delinking operation is symbolized by ‘\( = \)’.

\[
\begin{align*}
(13) & \quad \text{ROOT} \\
& \quad \text{+cons} \\
& \quad \text{+son} \\
& \quad \text{SUPRALAR} \\
& \quad [+\text{nas}] \\
& \quad \text{PLACE} \\
& \quad [...] \\
& \quad \text{ROOT} \\
& \quad \text{+cons} \\
& \quad \text{SUPRALAR} \\
& \quad \text{PLACE} \\
& \quad [...] 
\end{align*}
\]

14.4.1 Writing rules

It would be a very daunting prospect indeed if all rules were to be formulated by drawing these elaborate three-dimensional pictures. What we would need to do is draw just those elements in the configuration that are involved in the spreading and delinking, and make sure that the appropriate contextual conditions are
Distinctive features are phonetically non-abstract in the sense that they can be defined in terms of phonetic concepts. Would you say that natural feature classes are phonetically abstract?
**Q137** English has an optional rule assimilating [t d n] to the place of articulation of following velar and labial plosives and nasals, as shown in the data below. Formulate this English place assimilation in the display format.

| θɪn | θɪm bʊk | thin book |
| δæt | δæk kʌp | that cup |
| gʊd | gʊb bɔɪ | good boy |
| tɛn | tɛm mɑːlz | ten miles |
| δæt | δæt nɑt | that night |

**14.5 IMPLICATIONS OF UNDERSPECIFICATION**

Feature trees have led to more insightful formulations of phonological rules. Underspecification has helped to reach this result. We have just seen that not all segments have all nodes. Thus, laryngeal segments have no supralaryngeal node, and the place node has only those articulator nodes that positively specify the place of articulation of the segment, which implies, for instance, that [t] does not have a labial, dorsal or radical node (Sagey 1986; McCarthy 1988). In addition, predictable information can be considered absent in the underlying representation, being supplied by default rules at the end of the derivation. Below, a number of advantages of such underspecification of predictable surface features are briefly outlined.

**14.5.1 Default rules**

Leaving predictable features unspecified may have a simplifying effect on the formulation of rules that neutralize contrasts to some neutral, ’default’ realization. Let us take (Low) German final devoicing, which neutralizes the underlying forms /bʊnt/ ‘variegated’ and /bʊnd/ ‘union’ to [bʊnt], as an example (cf. the inflected forms [bʊntəs], [bʊndəs]). The expected laryngeal feature specification for obstruents is [–voice, –spread, –constr], which can be supplied as a default laryngeal node before the phonetic implementation rules start their work. Final devoicing can now consist of a rule delinking any underlying laryngeal node in coda position, as in (16).

```
(16) FINAL DEVOICING
    SC: ROOT
    SD: LARYNGEAL
    Coda
```
A further question is whether voiceless obstruents in a language like German could be left unspecified for the feature \([-\text{voice}]\) in the entire lexicon. Even though \([\pm\text{voice}]\) is contrastive for obstruents in German, which has contrasts like [pain] ‘pain’ – [bain] ‘leg’, it would be possible to leave voiceless obstruents unspecified, while specifying voiced obstruents as \([+\text{voice}]\). The voiceless obstruents then only acquire their \([-\text{voice}]\) specification at the end of the derivation, or, if the feature needs to be referred to by some rule, as early as it is needed. This is the assumption of Radical Underspecification (Kiparsky 1982a, 1993; Archangeli 1988), which goes one step further than the theory of Contrastive Underspecification discussed in section 6.3. A survey of the issues is provided by Roca (1994: ch. 2) and Steriade (1995).

Q138 Two northern dialects of British English each have a postlexical assimilation rule affecting consonant sequences across word boundaries (Wells 1982; Kerswill 1987). Neither rule occurs in the standard accent (Received Pronunciation (RP)).

1. Give SPE formulations of the two rules.
2. Explain why it would not be possible to assume that a default rule supplies the feature \([-\text{voice}]\) to English voiceless obstruents at the end of the derivation.

<table>
<thead>
<tr>
<th>English</th>
<th>RP</th>
<th>Yorkshire</th>
<th>Durham</th>
</tr>
</thead>
<tbody>
<tr>
<td>white blouse</td>
<td>[tb]</td>
<td>[tb]</td>
<td>[db]</td>
</tr>
<tr>
<td>wide shot</td>
<td>[df]</td>
<td>[tf]</td>
<td>[df]</td>
</tr>
<tr>
<td>ripe beans</td>
<td>[pb]</td>
<td>[pb]</td>
<td>[bb]</td>
</tr>
<tr>
<td>drab conditions</td>
<td>[bk]</td>
<td>[pk]</td>
<td>[bk]</td>
</tr>
<tr>
<td>black velvet</td>
<td>[kv]</td>
<td>[kv]</td>
<td>[gv]</td>
</tr>
<tr>
<td>five votes</td>
<td>[vv]</td>
<td>[vv]</td>
<td>[vv]</td>
</tr>
<tr>
<td>rough boys</td>
<td>[fb]</td>
<td>[fb]</td>
<td>[vb]</td>
</tr>
<tr>
<td>this village</td>
<td>[sv]</td>
<td>[sv]</td>
<td>[sv]</td>
</tr>
<tr>
<td>bad joke</td>
<td>[ddʒ]</td>
<td>[ddʒ]</td>
<td>[ddʒ]</td>
</tr>
<tr>
<td>live performance</td>
<td>[vp]</td>
<td>[fp]</td>
<td>[vp]</td>
</tr>
<tr>
<td>Bradford</td>
<td>[df]</td>
<td>[tf]</td>
<td>[df]</td>
</tr>
<tr>
<td>that night</td>
<td>[tn]</td>
<td>[tn]</td>
<td>[dn]</td>
</tr>
<tr>
<td>at last</td>
<td>[tl]</td>
<td>[tl]</td>
<td>[dl]</td>
</tr>
<tr>
<td>all true</td>
<td>[lt]</td>
<td>[lt]</td>
<td>[lt]</td>
</tr>
<tr>
<td>in Spain</td>
<td>[ns]</td>
<td>[ns]</td>
<td>[ns]</td>
</tr>
</tbody>
</table>

Q139 In Klamath, the underlying consonant pairs listed in the first column are pronounced as in the second column (cf. the data in (11); after Halle and Clements 1983). Two ordered rules are sufficient to derive
the outputs correctly. Assume that voiced sonorant consonants are unspecified for [+_voice], and that a default rule supplies this feature to unspecified sonorants at the end of the derivation. Also assume that if a geminate arises as a result of a spreading operation, the two original feature trees merge under a single ROuT node.

<table>
<thead>
<tr>
<th>Underlying</th>
<th>Derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>nl</td>
<td>ll</td>
</tr>
<tr>
<td>nl̥</td>
<td>lh</td>
</tr>
<tr>
<td>nlˀ</td>
<td>lʔ</td>
</tr>
<tr>
<td>ll̥</td>
<td>lh</td>
</tr>
<tr>
<td>llˀ</td>
<td>lʔ</td>
</tr>
</tbody>
</table>

1. State in prose what these two rules should accomplish.
2. Formulate the rules in autosegmental display notation. (Hint: Draw partial trees of a typical sequence like [nlˀ] before and after the change. Also draw a feature tree of [ll] after the merger under a single ROuT node.)

Q140 In Brao-Krung, voiced, voiceless unaspirated, voiceless aspirated and preglottalized voiced plosives are contrastive in the onset, as illustrated by [dak] ‘walk (verb)’; [truː] ‘fish’, [tʰun] ‘season’, [ʼduːr] ‘type of flute’ (the last may vary with an implosive [d]). In the coda, the only plosives that can occur are voiceless unaspirated, as illustrated by [neːt] ‘drink (verb)’ (Keller 2001).

1. Provide the full laryngeal feature specifications for the four types of initial plosive.
2. Which features could be supplied by default rules, assuming Radical Underspecification?
3. What underlying laryngeal feature specification would you assume for the coda plosives?

14.5.2 Long-distance assimilation

Underspecification equally aids the expression of long-distance assimilations. Articulator nodes frequently spread to the dominating node of an apparently non-adjacent segment. If the No CROSSING CONSTRAINT introduced in chapter 10 is valid in the segment tree, which of course it is, then this means that the intervening segment or segments must lack the dominating node in question. As an example, consider the Russian data in (17) (Kiparsky 1985). Sequences of obstruents in Russian, whether or not separated by a word boundary, agree in voicing with the right-most one. This is shown in (17a). As shown in (17b), however, it does not appear
to be the case that the obstruents must be strictly adjacent: a sonorant consonant may intervene.

\[
(17) \begin{array}{ccc}
\text{a} & \text{gorod-k-a} & \text{gorotka} & \text{‘little town (GEN)’} \\
\text{mtsensk bil} & \text{mtsenzg bil} & \text{‘it was Mtsensk’} \\
\text{b} & \text{iz mtsensk-a} & \text{is mtsenska} & \text{‘from Mtsensk’} \\
\text{ot mzd-i} & \text{od mzd} & \text{‘from the bribe’} \\
\end{array}
\]

Russian devoices final obstruents, a rule that is assumed to apply before the assimilation of voice illustrated above. The spreading feature can accordingly be assumed to be \([+\text{voice}]\). RUSSIAN VOICING ASSIMILATION is given in (18). The feature \([+\text{voice}]\) spreads left and any original specification for \([\text{voice}]\) is lost through delinking. This spreading may take place more than once, to deal with sequences of obstruents.

\[
(18) \text{RUSSIAN VOICING ASSIMILATION}
\]

However, the spreading of \([+\text{voice}]\) from the rightmost obstruent to a non-adjacent obstruent can only proceed if sonorants do not have a specification for \([\pm\text{voice}]\) in the underlying representation, just as \([\pm\text{constricted glottis}]\) and \([\pm\text{spread glottis}]\) are absent. Russian does not in fact contrast voiced and voiceless sonorants. In (19), we assume that the LARYNGEAL node is absent, though given as a bullet to identify the segment. The assimilation goes right through an intervening sonorant consonant, and the subsequent specification of the feature \([+\text{voice}]\) in sonorants is achieved by a default rule.

\[
(19)
\]

Again, given the absence of supralaryngeal features in glottal consonants, we can safely spread SUPRALARYNGEAL nodes across \([h \; \hat{\imath}]\). The fact that nasalization processes frequently spread through laryngeal segments to vowels in adjacent syllables thus finds a ready explanation under this assumption. Another oft-cited case of long-distance spreading relies on the absence of a CORONAL node in noncoronal (labial and dorsal) segments. Sanskrit retroflexion turns alveolar \([n]\) into retroflex \([\eta]\) after \([t \; s \; n \; \hat{\imath}]\). Retroflex consonants are contrastively \([-\text{ant}, -\text{distr}]\), as shown in (20) (Steriade 1987; Kenstowicz 1994a).
We would at first sight expect the rule to be blocked by any consonant other than [n]. However, noncoronal consonants as well as vowels may freely intervene between target and trigger, as shown in the first column of (21). By contrast, there may not be a coronal consonant other than [n] between target and trigger, as shown in the second column of (21). (Underlining voiceless obstruents are shown as voiced between sonorant segments.)

The rule to accomplish this is given in (22), which relies crucially on the absence of a coronal node between the two segments. If there is a coronal node, the nasal cannot be reached, as the spreading association line would cross the association line between the coronal node and the place node, as shown in (23).
Q142 Stem-initial obstruents in Bakairí, a language whose syllable structure is (C)V, are realized as voiceless plosives, unless preceded by a vowel, when they are voiced (Wetzels 1997b). This is shown by [ɔkɔ] ‘bow’, [tɔ-dɔka-ge] ‘have a bow’. A stem-internal obstruent, however, can be voiced or voiceless, as shown by [itubi] ‘skin’, [ɔdɔpigi] ‘heat’.

1 Assume that stem-initial obstruents are unspecified for [±voice] in underlying representations. Which two default rules must be assumed?

The distribution of stem-internal obstruents is not free. If stem-initial consonants are not counted, there can be only one voiceless obstruent in the stem, which must be either the first or the second obstruent (again, not counting any stem-initial consonant). Thus, while [tɔkɔ] ‘bow’, [ɔdɔpigi] ‘heat’ and [ɔdɔdɔ] ‘jaguar’ are fine, forms like *[itupi], *[t̥itupi], *[idebiko] are impossible. This state of affairs can be captured by assuming that Bakairí stems have a limited number of [±voice] melodies, much in the way that Etung has a limited number of tone melodies.

2 Assuming that [-son] segments are the [±voice]-bearing segments, which [±voice] melodies would you assume? Would the feature spread?

3 How are [ɔdɔdɔ] and [ɔdɔpigi] accounted for?

4 Give the underlying forms of [tɔkɔ] ‘bow’ and [tɔdɔkage] ‘have a bow’, indicating obstruents unspecified for voice with capital letters P, T, K. Illustrate the association of [±voice] melodies and the effects of the default rules.

5 In [kɔnɔpĩo] ‘little bird’, the voiceless [p] is the third consonant of the stem. Why isn’t this form ungrammatical?
14.5.3 Vowel harmony

Vowel harmony is a subclass of long-distance assimilation. Many languages exclude certain combinations of vowels in the word (for a survey, see van der Hulst and van de Weijer 1995). For instance, a Finnish word containing one of the back vowels [u o a] must not contain any of the front counterparts [y ø æ]. Thus, possible words are [mite] ‘what’, [suomi] ‘Finland’, [talo] ‘house’, while *[tymo] and *[tumä] are ill-formed. Since the vowels must agree for the feature [±back], Finnish represents a case of Back Harmony. Languages may also or additionally display Round Harmony, Height Harmony and ATR-Harmony. In Akan, the four [+ATR] vowels [i e o u] must not co-occur with any of their [–ATR] counterparts [i e o u]. This situation is again reminiscent of the restriction on tonal patterns in Etung and Mende discussed in chapter 10. Specifically, it would be reasonable to assume that, underlyingly, backness in Finnish is a property of the word, rather than of the individual vowels. Thus, underlyingly, the vowels of a Finnish word like [pysæhtyæ] ‘to stop’ would be unspecified for [±back]. Instead, the word appears in the lexicon with a ‘floating’ feature [–back] (24a), which associates with the dorsal nodes of all the vowels, as shown in (24b), which is equivalent to (24c). (In (24a,b) the underspecification for [±back] is indicated by the capital letters. Thus, [A] could be [æ] or [a], etc.) Notice that there is little point in assuming a particular direction of association, as there is only one feature to be associated. The important advantage of the lexical representation in (24a) is that the mutual exclusion of [+back] and [–back] vowels in the same word is naturally accounted for.

\[
\begin{array}{ccc}
(24) & a & pUsAhtUA \\
& b & pUsAhtUA \\
& & [–back] \\
& & \downarrow \\
& & c & [pysæhtyæ] \\
& & [–back] \\
\end{array}
\]

The vowel systems of languages with vowel harmony frequently have one or more vowels which do not contrast for the harmony feature of the language. Thus, in addition to the two sets of four vowels mentioned above, Akan has a [–ATR] [a], for which there is no [+ATR] counterpart *[a]. Similarly, Finnish has the two [–back] vowels [i e], for which there are no [+back] counterparts *[u y]. Depending on the language, such noncontrasting vowels behave in either of two ways. They may be opaque, in which case they stop the spreading of the feature with the opposite value and, moreover, impose their value on any following vowels. Alternatively, they are transparent, in which case they allow the spreading of the harmony feature to go right through them. Opaque vowels thus block and impose harmony, while transparent vowels neither undergo nor impose it (Gafos and Dye 2011). Akan [a] is an opaque vowel, while Finnish [i e] are transparent vowels.

As was shown by Clements (1981), the autosegmental model has no problem in characterizing the behaviour of opaque vowels. If we assume that the floating feature only spreads to vowels that are unspecified for the harmony feature, we predict that the spreading is stopped by any vowel that already has a specification. Akan words are made up of a root, which has either floating [+ATR] or floating [–ATR], plus a number of prefixes and suffixes. If the word contains only nonlow vowels, all
the vowels will agree with the feature value of the root, as illustrated by [e-bu-o] ‘nest’, which contains the [+ATR] root [bu], and [e-bu-ɔ] ‘stone’, which contains the [−ATR] root [bu]. However, [a] may occur in combination with either [+ATR] or [−ATR] vowels, as shown by the disyllabic roots [pîrə] ‘to sweep’, [järɪ] ‘to be sick’, [bîsə] ‘to ask’ and [kərɪ] ‘to weigh’, where the last two are ‘disharmonic’ (Clements 1981). Their lexical representations are given in (25). (The vowel [a] has a number of allophones, which we have not indicated.)

\[(25)\]

\[\begin{array}{cccc}
 a & pɪrə & b & järɪ \\
 & [−ATR] & & [−ATR] \\
 & & c & bîsə & d & kərɪ \\
\end{array}\]

Disharmonic roots will have prefixes and suffixes that surface with opposite values for [ATR]: the floating [+ATR] feature and the associated feature of [a] will each spread to their ‘half’ of the word, as shown in (26), where (26a,b) are equivalent to (26c,d), respectively.

\[(26)\]

\[\begin{array}{ccc}
 a & 0 - bîsə - 1 & b & 0 - kərɪ - 1 \\
 & [+ATR][−ATR] & & [−ATR][+ATR] \\
 c & [o - bîsə - 1] & d & [o - kərɪ - 1] \\
\end{array}\]

Q143  Akan does not have words with [+ATR] vowels on both sides of [a]. Does the above account explain this?

By contrast, the behaviour of transparent vowels has been problematic for the autosegmental model. It would be reasonable to want to represent transparent vowels as underspecified for the harmony feature of the language, so that they can let the spreading feature through. However, since harmonizing vowels are crucially unspecified for the harmonizing feature, a Back Harmony rule, for instance, cannot be expected to distinguish between harmonizing and transparent vowels during its search for empty dorsal nodes in vowels. There is no hope for a general solution based on leaving the entire dorsal node out in the case of transparent vowels. This will work in the case of languages that have only one transparent vowel, but not in the case of Finnish, which has two transparent vowels, [i] and [e], which contrast for a feature which in the consensus model is present on the dorsal node: [±high]. The solution here must be an appeal to a constraint banning ill-formed vowels. For Finnish, the constraint would ban [ui] and [y], i.e. *[−round, +back, −low]. The existence of transparent vowels may thus serve as a further argument for abandoning rule-based descriptions in favour of constraint-based ones.
Q144 The Finnish word game Siansaksa ('Pig German') produces the forms in the second column for the Finnish forms in the first (Vago 1988).

a  saksalaisia hætyːtetːiːn  hæksælæisiæ satuːtetːiːn
   ‘the Germans were attacked’

b  tykːæːn urheilusta  ukːaːn tyrheilystæ
   ‘I like sports’

c  otsansa hiesːæ  hitsansa oesːa
   ‘in the sweat of his brow’

d  pitæː kalasta  kataː pilasta
   ‘likes fish’

1  Give an analysis of Siansaksa.
2  Does your analysis account for the fact that the first word in the game form in (a) is not *[hæsælæisiæ]?  
3  Does your analysis account for the fact that the second form in the game form in (a) is not *[satyːtetːiːn]?  
4  Does your analysis account for the fact that in (c) the first word in the game form is not *[hitsænsæ]?  

14.6 COMPLEX SEGMENTS

Unlike simplex segments, complex segments have more than one specification either for place of articulation or a manner feature. Three types of complex segments can be distinguished.

1  Complex-place segments. A complex-place segment has more than one articulator node, and more than one articulator therefore participates in realizing the constriction specified by the features in the root. That is, a double articulation is produced. Examples are [gb kp ŋm], all of which are labial-dorsal. English [w] is a labial-dorsal approximant. Labial-coronal [pt] occurs in Bura and Margi. In (27), a labial-coronal segment is depicted. Of course, above the place node the segment looks just like a simplex segment. This is an end view: the two articulator nodes lie in the segment plane, and are thus phonologically simultaneous.

(27) PLACE
   LABIAL  CORONAL
2 **Manner-contour segments.** Manner-contour segments have a sequence of differently valued occurrences of the same manner feature. For example, prenasalized stops like [mb nd ng] are [+nas] as well as [–nas]. Since [±nasal] defines a feature plane, this can only be represented by arranging the two specifications in sequence, and a side view of such a segment would look like (28).

![Diagram of manner-contour segment](image)

A second type of manner-contour segment has been claimed by Sagey (1986), among others, for affricates, like German [ts] as in [tsu:] ‘closed’ or English [tʃ] as in [tʃu:] *chew*. In this view, affricates have a sequence of [–cont][+cont], instead of being simplex segments that are characterized by [+strid], to distinguish them from [–strid] plosives. The most important argument for the view that affricates are complex is that they are *pronounced* as complex segments, a closure followed by a narrowing that causes friction. However, there is evidence that affricates function as strident plosives and are simple [–cont] segments. If they were (also) specified as [+cont], they would be expected to pattern with fricatives, but this appears not to be the case. For instance, Polish [s] is assimilated to [c] by a following [tɛ], as in [œtɛcɛ] ‘thistle+SG, LOCATIVE’, from [œstɛcɛ], but the affricate [ts] is unaffected, as in [œtstɛcɛ] ‘vaccine+SG;LOCATIVE, *[œtɛcɛ]’ (Rubach 1994; Kim 2001). Of course, under either analysis, an affricate is a single segment, but the evidence appears to be in favour of a simplex single segment (see also Q47 and Berns 2013).

3 **Secondary articulations.** Consonants with a secondary articulation have two place specifications: one to indicate the location of the manner of articulation, and one to indicate a simultaneous vocalic articulation. That is, while the segment is [+cons], and can have any configuration of manner features, it has an additional component specifying a vowel-like gesture of either the lips or the tongue body. A secondary articulation like labialization can combine with a labial, a coronal or a dorsal segment. The labialized segments [pʰ wʰ tʰ nʰ kʰ ɲʰ] are only some of the contrastively labialized segments of Nambakaengo. Similar freedom of combination can be shown to exist for palatalization, velarization and pharyngealization. There is, however, a strong tendency not to combine identical place specifications. While segments like [χʰ], [χʰ] are relatively common, labialized labials are rare. (Velarized velars have not been reported; indeed, it is not clear how such segments could be different from plain velars.)
The representation of secondary-articulation segments involves the addition of a **PLACE** node. However, unlike complex-place segments, the two place specifications of secondary-articulation segments differ in manner of articulation. In order to indicate which of the place features specifies the secondary articulation and which the consonantal place, one or the other must be marked with some diacritic, as in (29), which shows a partial representation of \( [p'] \).^2

\[
(29) \quad \text{SUPRALAR} \quad \left[ \text{–cont} \right] \quad \text{PLACE} \quad \text{LABIAL} \quad \text{DORSAL}_{\text{secondary}} \quad \left[ \text{–back} \right]
\]

The two **PLACE** nodes in a segment with secondary articulation are not sequenced in time. Although in IPA the superscripts indicating labialization, velarization, etc. conventionally appear to the right of the consonant symbol, the two components of a secondary-articulation segment are phonologically simultaneous. That is, a side view would show a straight line.

**Q145** There have been no reports of complex segments like \( [px] \) or \( [dn] \). How do the representations assumed above account for this?

### 14.6.1 Evidence for complex segments

Like single segments, complex segments have a single **ROOT** node, and as such distinguish themselves from sequences of segments. This structural difference is reflected in their behaviour in a number of ways. An obvious place to look for such single-segment behaviour is the way they are incorporated into syllable structure. At first sight, we might be tempted always to regard phonetic sequences like \( [mb] \) or \( [nd] \) as sequences of phonological segments. However, the fact that in Bakwiri, for example, these are the only phonetic sequences ever to appear prevocally is a strong indication that \( [mb] \), \( [nd] \), \( [ŋg] \) are single segments in this language, as this analysis will allow us to restrict its syllable structure to CV. It explains why no other combinations of consonants ever occur. Under the alternative analysis with CC-onsets, an explanation would have to be found for the peculiar restriction that the only complex onsets consist of nasals followed by plosives with the same place of articulation, even though complex onsets like
[st-] or [pj-] are common in other languages. Another alternative that must be rejected is an assignment of the nasal and the plosive to different syllables, assuming CV(C) as the syllable structure. In that analysis, [limba] would be syllabified as [lim.ba] rather than [lim.bla], while a word like [mbila] would have to be [m.bi.la], with a syllabic [m]. This analysis can be empirically excluded on the basis of the Bakwiri word game that appeared in Q96, in which the syllables of disyllabic words are reversed and [kóndi] ‘rice’ is transformed to [ndíkò] (Hombert 1973, 1986). Clearly, if the language were to allow codas, we would expect the syllable structure of the word to be *[kó.ndí], and the inverted form to be *[díkòn]. This, however, is not the correct game form.

Q146 In Ewe, verb stems reduplicate to form a present participle, noun or adjective. Determine what segments of the verb stem are reduplicated. Why do these data (from Sagey 1986: 86) demonstrate that Ewe has compound-place segments and an affricate?

<table>
<thead>
<tr>
<th>Stem</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>fo</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>zo</td>
<td>‘to walk’</td>
</tr>
<tr>
<td>aha+no</td>
<td>‘liquor + drink’</td>
</tr>
<tr>
<td>fja</td>
<td>‘to burn’</td>
</tr>
<tr>
<td>bja</td>
<td>‘to ask’</td>
</tr>
<tr>
<td>fle</td>
<td>‘to buy’</td>
</tr>
<tr>
<td>npra</td>
<td>‘to rave’</td>
</tr>
<tr>
<td>kplo</td>
<td>‘to lead’</td>
</tr>
<tr>
<td>avosja</td>
<td>‘cloth + to dry’</td>
</tr>
<tr>
<td>tsi</td>
<td>‘to grow’</td>
</tr>
<tr>
<td>gbla</td>
<td>‘to exert oneself’</td>
</tr>
</tbody>
</table>

14.6.2 The representation of palatals and palatoalveolars

In the consensus model (Broe 1992), roundness requires a labial articulator node, and tongue body specification requires a dorsal articulator node; in this account, round vowels have two articulator nodes. Segments like [y ø u] are therefore complex, i.e. LABIAL-DORSAL. Recall from Q145 that the place specification of the segment [u] is identical to that of the LABIAL-DORSAL approximant [w], as well as to that of the LABIAL-DORSAL plosive [gb]. But while the assumptions made so far correctly characterize [u] and [w] as a natural class, they fail to characterize [i] and [j] as having the same place of articulation. The assumption that [i j] and [u w] form natural classes is supported by processes in a number of languages which have alternations between [i] and [j] as well as between [u] and [w]. In such cases,
[i u] appear in the syllable peak, while [j w] appear elsewhere. Our treatment so far has assigned the feature [CORONAL] to palatoalveolars and (pre)palatalis like [j], but [DORSAL] to vowels, including front vowels like [i]. Not only does this make it impossible to describe alternations between [i] and [j] as involving a difference in position in syllable structure, it also fails to account for a number of effects [i] and [j] have in common (Bhat 1978; Lahiri and Evers 1991).

1 The segment [j] triggers fronting of velars, just like [i] (and possibly other front vowels). For instance, French [kjoʃk] ‘kiosk’ has the same [k] as does [ki] ‘who’. The fronting is easily explained for [i], which is [DORSAL, –back]. But why should [j], which is [CORONAL, –ant, +dist], have the same effect?

2 In many languages, [k] before [i] has changed into prepalatal (i.e. [CORONAL]) [c], and from there into palatoalveolar [tj]. Taking this route, Latin [ki:wi‘ta:tem] and [‘kirkə:] ended up as [tʃit‘ta] ‘city’ and [‘tʃirka] ‘approximately’ in Italian. If [i] is [DORSAL], how can it affect a dorsal consonant so as to become coronal?

3 Many languages, of which Japanese is an example, have processes that change [ts] into [te c] before [i], as in [sitimi], pronounced [cicimical], ‘pepper’. [te c] are coronal. If [i] is [DORSAL], then why don’t we get dorsal [çi kimì] instead?

Beginning with Keating (1988), many researchers have analysed segments like [j] as compound-place segments with both a coronal and a dorsal articulator. In this view, [j i] both have the structure as in (30). With Jacobs (1989) and Jacobs and van de Weijer (1992), we assume that palatoalveolars are also CORONAL-DORSAL. This view is supported by articulatory data, in the sense that all these segments involve articulations between the forward part of the tongue and a section of the palate stretching some 2 cm back from the alveolar ridge. These data have been obtained with the help of X-ray pictures with side views of the tongue, as well as with the help of electropalatography, a technique for recording contact areas of the roof of the mouth (Recasens et al. 1995).

(30)
This feature analysis accounts for all three earlier observations. First, French velar fronting of \([k]\) to \([k^\circ]\) in the environment of \([j\ i]\) can be accounted for as spreading of the \([\pm \text{back}]\) node of \([j]\) and front vowels to the dorsal node of the velar. This is shown in (31), where the \([-\text{back}]\) feature of \([j\ i]\) will associate with the dorsal node of \([k]\). (Depending on the language, the rule may have to be constrained to prevent other front vowels from triggering it.)

(31) VELAR FRONTING

This analysis also accounts for the occurrence of \([c]\) instead of velar \([x]\) in the environment of \([j]\) and front vowels, as in Greek \(\acute{\text{cilia\}}\) 'thousand' versus \(\acute{xo\}'\)ros 'dance'. That is, the fronting of \([x]\) to \([c]\) amounts to the replacement of \([+\text{back}]\) with \([-\text{back}]\) in the dorsal fricative.3

The assimilation of \([k]\) to \([c]\) or \([t\f]\) can be described as the spreading of the coronal-dorsal node of \([j\ i]\) to the supralar node of the velar, as shown in (32), where the left-hand segment, \([k\ g]\), receives the place node of the following \([i\ j]\).

(32)

And, to complete the story, (33) describes the assimilation of \([t\ s]\) to \([t\ c]\) or \([t\f\ s]\) as the spreading of the place node from \([j\ i]\) to the consonant. This analysis predicts that the coronal specification of \([t\ d\ s\ z]\) becomes \([-\text{ant}]\), which is correct. On the distinction between prepalatal \([c]\) and palatoalveolar \([\j]\), see the next section. Note that these spreading operations only account for the place assimilation. The manner features will in many cases be supplied separately. For instance, if the product of (33) is \([c]\), nothing needs to be done, but if it is \([t\f]\), \([+\text{strid}]\) needs to be added (cf. Kim 2001).

(33)
Q147  Standard Chinese contrasts three coronal fricatives [s § ɕ] in the onset, as shown in (a,b,c). Before [i], only [ɕ] occurs; after [s §], syllabic approximants occur which are homorganic with the fricatives, and which historically go back to [i] (Lee-Kim 2014).

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[sa\textsuperscript{55}]</td>
<td>‘spill’</td>
</tr>
<tr>
<td>d</td>
<td>[si\textsuperscript{55}]</td>
<td>‘tear (Verb)’</td>
</tr>
<tr>
<td>b</td>
<td>[sa\textsuperscript{55}]</td>
<td>‘murder’</td>
</tr>
<tr>
<td>e</td>
<td>[ʂ\textsuperscript{55}]</td>
<td>‘wet’</td>
</tr>
<tr>
<td>c</td>
<td>[ɕa\textsuperscript{55}]</td>
<td>‘foolishly’</td>
</tr>
<tr>
<td>f</td>
<td>[ɕi\textsuperscript{55}]</td>
<td>‘West’</td>
</tr>
</tbody>
</table>

1. What is the contrastive specification under [CORONAL] for the three fricatives (cf. section 6.3)?
2. If the V-position has a segmental node with [−cons, +son], can the approximant consonant be described by the spreading of a single node?
3. While it is clear that the approximant consonant is in the nucleus of the syllable, the question may arise whether this segment is a vowel or a syllabic consonant. Assuming that this distinction is meaningful, can it be characterized as a consonant in our feature model?
4. What does this assumption imply for the contrast between [i] and [j]?
5. What does this assumption imply for the contrast between [u] and [w]?

14.6.3  Distinguishing among palatoalveolars

The coronal-dorsal representation of [ʃ ʒ ʨ ʥ] would appear to offer a satisfactory description of these consonants, which are typically pronounced with the tongue tip raised towards, or articulating with, the alveolar ridge and a simultaneous raising of the tongue blade and forward part of the front towards the postalveolar area of the hard palate. The CORONAL node is responsible for specifying the first type of articulation, while the DORSAL node is responsible for the second aspect. Of course, the CORONAL node can dominate a number of specifications for the features [±anterior] and [±distributed]. Even though different palatoalveolar consonants rarely contrast in the same language, the implicit prediction that there are different types of palatoalveolar consonants is certainly correct. First, a [−ant, +distr] articulation will have the blade articulating with the postalveolar region, while allowing the tongue tip to be behind the lower teeth. This is typically the tongue position for [j] and [i]. In the case of [±cons] segments, this type is sometimes referred to as ‘prepalatal’ or ‘alveopalatal’, symbolized [c z ʨ ʥ]. The fricatives and affricates may contrast with other palatoalveolars. In Serbo-Croat, for instance, [ʨ] contrasts with [ʦ], for which the coronal contact is postalveolar or retroflex, i.e. [−ant, −distr], as shown by [ʦ̑ːr] ‘enchantment’ vs [ʨ̑ːr] ‘profit’. Both CORONAL-DORSAL affricates are distinct from the coronal affricate [ʦ], as shown by [ʦ̑ːːriːti] ‘to leak’ vs [ʨ̑ːːriːti] ‘to blow (of wind)’ and [ʦ̑ːh] ‘gild, union’ vs [ʨ̑ːh] ‘Czech man’. Similarly, Polish contrasts retroflex [ʂ ʈ] with alveopalatal [ɕ ʨ] as in prosze, [prɔʂɛ] ‘please’, czas [tʃas] ‘time’ vs prosie, [prɔɕɛ] ‘piglet’ and ćiasny [tʃasni] ‘tight’. 
English [ʃ tʃ dӡ] are [+ant, –distr]: in addition to the raised front and/or blade, there is an apical articulation with the alveolar ridge. This representation accounts for an apparent asymmetry in the application of coronal assimilation in English. Recall that this process causes coronal [t d n l] to adopt the place of articulation of a following coronal consonant, so that the [t] of that is dental in that thing, alveolar in that tin and postalveolar in that trip (cf. section 14.2). The apparent asymmetry is that before palatoalveolars the place of articulation of [t d n l] is unaffected. Even though the forward part of the front may show some anticipatory raising towards the hard palate, [t d n l] retain their contact between the tip and the alveolar ridge, just as they do when a vowel follows, so that the location of the crown during [t] in that egg is the same as that during [t] in that chore, and the [n] in keen officer is no different from the [n] in keen judge.4 Our analysis readily accounts for these data: [tʃ dʒ] have a coronal node with [+ant, –distr], which explains the articulatory action of the tongue tip, in addition to a dorsal node with [–back], which explains the raising of the front of the tongue. Since it is only the coronal node that spreads to a preceding [t d n l], the effect on their place of articulation is entirely vacuous, because these consonants are already [+ant, –distr]. By contrast, if we were to analyse palatoalveolars as simplex coronal segments, as was done in chapter 6, they would have to be [–ant, +distr], a featural characterization that would be necessary to explain the largely postalveolar place of articulation. This would incorrectly predict that coronal assimilation causes a considerable retraction of the place of articulation of [t d n l] before palatoalveolars. We summarize our feature analysis of coronal, coronal-dorsal and dorsal consonants in Table 14.1.

### Table 14.1
Feature analysis of coronal, coronal-dorsal and dorsal consonants. ✓ stands for the presence of a unary feature.

<table>
<thead>
<tr>
<th></th>
<th>CORONAL</th>
<th>CORONAL-DORSAL</th>
<th>DORSAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apical [t s]</td>
<td>Prepalatal [i j ɹ e]</td>
<td>Eng. [ʃ tʃ dӡ]</td>
</tr>
<tr>
<td>Coronal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Anterior</td>
<td>+</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Distributed</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Dorsal</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Back</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Q148 Tahltan has the following system of oral coronal [+cons] segments (Shaw 1991):

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>tɬ</td>
<td>tθ</td>
<td>ts</td>
<td>tʃ</td>
</tr>
<tr>
<td>d</td>
<td>dl</td>
<td>ð</td>
<td>dz</td>
<td>dz</td>
</tr>
<tr>
<td>t’</td>
<td>tɬ’</td>
<td>tθ’</td>
<td>ts’</td>
<td>tʃ’</td>
</tr>
<tr>
<td>l</td>
<td>θ</td>
<td>s</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>δ</td>
<td>z</td>
<td>ɹ</td>
<td></td>
</tr>
</tbody>
</table>
The dental, alveolar and palatoalveolar consonants in the last three columns are involved in **CORONAL HARMONY**. Within the word, the rightmost consonant in this group determines the place of articulation of all preceding consonants in the same group. This is illustrated with [s] ‘first person singular subject’, and with [θi] ‘first person dual subject’ in the following data. The (a) examples show these suffixes in their unassimilated form, while the (b) examples show the effect of **CORONAL HARMONY**. The consonants in the first two columns above are transparent to **CORONAL HARMONY**: they neither trigger nor block the process. This is illustrated by the (c) examples. Target consonants are underlined.

![Feature geometry](257)

<table>
<thead>
<tr>
<th>Underlying [s]</th>
<th>Underlying [θi]</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
</tr>
<tr>
<td>εšk`a:</td>
<td>‘I’m gutting fish’</td>
</tr>
<tr>
<td>εšdan</td>
<td>‘I’m drinking’</td>
</tr>
<tr>
<td>neštel</td>
<td>‘I’m sleepy’</td>
</tr>
<tr>
<td>hudi[ʃ]a</td>
<td>‘I love them’</td>
</tr>
<tr>
<td>lene[ʃ]{u}:</td>
<td>‘I’m folding it’</td>
</tr>
<tr>
<td>εθdù:θ</td>
<td>‘I whipped him’</td>
</tr>
<tr>
<td>meθeq</td>
<td>‘I’m wearing’</td>
</tr>
<tr>
<td>ededəθdù:θ</td>
<td>‘I whipped myself’</td>
</tr>
<tr>
<td>jaʃ[t]:tʃ</td>
<td>‘I splashed it’</td>
</tr>
</tbody>
</table>

1. How can we tell from these data that the two suffixes have different underlying forms?
2. What are the three place specifications of the consonants involved in **CORONAL HARMONY**?
3. Why is it impossible to view **CORONAL HARMONY** as the leftward spreading of a [coronal] node to consonants that are unspecified for [coronal]?
4. Is it possible to view **CORONAL HARMONY** as a rule that spreads a [coronal] node to consonants that have a [coronal] node, while skipping consonants that do not? Motivate your answer.

### 14.7 CONCLUSION

Many languages place restrictions on the combination of different values of the same phonological features in the word. Such word-based distributional restrictions include vowel harmony and consonant harmony, as well as word-based distributional patterns of nasalization and obstruent voicing. The autosegmentalization of features, together with the underspecification of certain features in certain segments, allows us to describe these patterns with the help of many-to-one associations, a more insightful solution than is possible if segments are represented
as unstructured lists of features, as was usual in the linear theory of SPE. To fully account for transparent segments, however, a rule-based description will need to fall back on constraints. This can be seen as an argument for abandoning the concept of rules that literally change one representation into another, and for adopting a theory like Optimality Theory, which relies on constraints that simply tell us what surface representations must look like. We also saw that the feature tree allows for a natural representation of complex segments. It was shown that, in addition to obviously complex segments like [mb], [kp] and [tʰ], palatal and palatoalveolar consonants can fruitfully be interpreted as (complex) CORONAL-DORSAL segments, a type of representation that makes it possible to describe various place assimilations involving palatoalveolar articulations.

In this book, we have seen an impressive array of the ways in which languages organize the sound resources that are available to them. We have seen how phonological theory has responded to this variation by constructing a representational framework in which phonological content is slotted into a hierarchically organized array of structural positions. These structures and the nature of the phonological content in them intrigue us because of the shifting perspective they offer between how similar they are across languages and how extraordinarily different. Phonological theory has been quite successful in laying out these representations. Yet we have also seen how theoretical accounts of the relations between different levels of representation are not fully satisfactory. Such issues are likely to be addressed in behavioural and neurological psycholinguistic research. Equally, we may look forward to insights that a continued expansion of the data base will bring through perceptive analyses of undescribed and underdescribed languages.

NOTES

1 To cater for the two kinds of evidence, Lombardi (1990) proposed an underlying representation with an unordered [–cont]/[+cont] specification, which become ordered on the surface.

2 In Clements and Hume (1995), the feature tree includes a VOCALIC PLACE node in addition to a CONSONANTAL PLACE node, each of which dominates the usual articulator nodes. A feature like CORONAL is interpreted as a front articulation in the case of vowels and as a coronal pronunciation in the case of consonants, while DORSAL is used to characterize back vowels or dorsal consonants. A separate node is used to specify the degree of opening of vowels. See also Kenstowicz (1994a).

3 The [ç] is classed as a complex coronal-dorsal by Keating (1988). In addition to making it impossible for [ʃ] to be coronal-dorsal, this characterization does not do justice to the more retracted articulatory contact for [ç] than for the palatoalveolars and [j]. Also, the implied prediction that the fronting of [k] is a different process from the fronting of [x] has not been confirmed.

4 Clements (1985) noted that the articulation of [t d n], but not of [l], changes before palatoalveolars. At least for standard British English, we believe there is no such distinction.
References

This list includes references to works in the Key to the Questions, available in the eResource at www.routledge.com/9781138961425.


Morais, José; Cary, Luz; Alegria, Jésus; Bertelson, Paul (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition* 7: 323–331.


References


References


Volk, Erez (2011). Mijikenda phonology. PhD, Tel Aviv University, Tel Aviv.


The language index below was modelled on the one in Laver (1994) and composed with the help of the Ethnologue (Lewis at al. 2016) and the Glottolog (www.glottolog.org).

<table>
<thead>
<tr>
<th>Language</th>
<th>Family</th>
<th>Geographical area</th>
<th>Page numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>!Xū</td>
<td>Khoisan</td>
<td>Angola, Namibia</td>
<td>4, 37, 41, 45</td>
</tr>
<tr>
<td>Akan</td>
<td>Niger-Congo, Kwa</td>
<td>Ghana (South)</td>
<td>78, 247</td>
</tr>
<tr>
<td>Amele</td>
<td>Trans-New Guinea, Gum</td>
<td>Papua New Guinea (Madang province)</td>
<td>182–3</td>
</tr>
<tr>
<td>Amharic</td>
<td>Semitic</td>
<td>Egypt, Ethiopia</td>
<td>85, 132</td>
</tr>
<tr>
<td>Angaathiha</td>
<td>Trans-New Guinea, Angan</td>
<td>Papua New Guinea (Morobe province)</td>
<td>39–40</td>
</tr>
<tr>
<td>Arabic</td>
<td>Semitic</td>
<td>Algeria, Bahrain, Chad, Egypt, Ethiopia, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates, Yemen</td>
<td>29, 33, 38, 46, 79, 130, 132, 149, 178</td>
</tr>
<tr>
<td>Cairene Arabic</td>
<td></td>
<td>Egypt (Cairo)</td>
<td>184</td>
</tr>
<tr>
<td>Gulf Arabic</td>
<td></td>
<td>Bahrain, Kuwait, Oman, Qatar, United Arab Emirates</td>
<td>29</td>
</tr>
<tr>
<td>Arawak</td>
<td>Arawakan</td>
<td>Surinam (Commewijne and Para districts)</td>
<td>31</td>
</tr>
<tr>
<td>Bakairi</td>
<td>Carib</td>
<td>Brazil (Mato Grosso state)</td>
<td>45, 246</td>
</tr>
<tr>
<td>Bakwiri</td>
<td>Niger-Congo, Bantu</td>
<td>Cameroon (South)</td>
<td>38, 157, 252</td>
</tr>
<tr>
<td>Balantak</td>
<td>Austronesian, Malayo-Polynesian</td>
<td>Indonesia (East Central Sulawesi)</td>
<td>105</td>
</tr>
<tr>
<td>Barasana</td>
<td>Tucanoan</td>
<td>Colombia (Amazonas department)</td>
<td>160–1, 163</td>
</tr>
<tr>
<td>Basque</td>
<td>Isolate</td>
<td>Spain (Basque country)</td>
<td>204</td>
</tr>
<tr>
<td>Bengali (aka Bangla)</td>
<td>Indo-European, Indo-Aryan</td>
<td>Bangladesh, India</td>
<td>77, 178–9, 204, 207</td>
</tr>
<tr>
<td>Brao-Krung</td>
<td>Austro-Asiatic, Brao-Kravet</td>
<td>Cambodia (Ratanakiri province), Laos</td>
<td>243</td>
</tr>
<tr>
<td>Bugotu</td>
<td>Austronesian, Oceanic</td>
<td>Solomon Islands (Isabel province)</td>
<td>140</td>
</tr>
<tr>
<td>Bura</td>
<td>Afro-Asiatic, Chadic</td>
<td>Nigeria (Borno state)</td>
<td>249</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------</td>
<td>--------------------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Burmese</td>
<td>Sino-Tibetan,</td>
<td>Myanmar</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>Tibeto-Burman</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cantonese</td>
<td>Sino-Tibetan, Chinese</td>
<td>China (South), Hong Kong</td>
<td>51</td>
</tr>
<tr>
<td>Catalan</td>
<td>Indo-European, Romance</td>
<td>Spain (Catalonia, Valenciana, Balearic Islands)</td>
<td>31, 39, 136, 175</td>
</tr>
<tr>
<td>Cavineña</td>
<td>Tacanan</td>
<td>Bolivia (El Beni department)</td>
<td>194, 197</td>
</tr>
<tr>
<td>Chama</td>
<td>Tacanan</td>
<td>Bolivia (El Beni department)</td>
<td>197</td>
</tr>
<tr>
<td>(aka Ese Eja)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chipewyan</td>
<td>Athabaskan</td>
<td>Canada (South Central)</td>
<td>32</td>
</tr>
<tr>
<td>(aka Dene)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chuave</td>
<td>Trans-New Guinea, Chimbu</td>
<td>Papua New Guinea (Chimbu province)</td>
<td>41, 46</td>
</tr>
<tr>
<td>Cilungu</td>
<td>Niger-Congo, Bantu</td>
<td>Zambia (Northeast)</td>
<td>164</td>
</tr>
<tr>
<td>Ciyao</td>
<td>Niger-Congo, Bantu</td>
<td>Malawi, Mozambique, Tanzania</td>
<td>154</td>
</tr>
<tr>
<td>Corsican</td>
<td>Indo-European, Romance</td>
<td>France (Corsica)</td>
<td>71, 81, 224</td>
</tr>
<tr>
<td>Creek</td>
<td>Muskogean</td>
<td>USA (Alabama; Oklahoma)</td>
<td>195</td>
</tr>
<tr>
<td>Czech</td>
<td>Indo-European, Slavic</td>
<td>Czech Republic, Slovakia</td>
<td>31, 39, 46, 255</td>
</tr>
<tr>
<td>Danish</td>
<td>Indo-European, Germanic</td>
<td>Denmark</td>
<td>81</td>
</tr>
<tr>
<td>Dinka</td>
<td>Nilo-Saharan</td>
<td>South Sudan</td>
<td>22–3, 39, 137, 165</td>
</tr>
<tr>
<td>Dizi</td>
<td>Afro-Asiatic</td>
<td>Ethiopia (Southwest)</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Netherlands Antilles, Surinam</td>
<td></td>
</tr>
<tr>
<td>Kaatsheuvel</td>
<td>Dutch</td>
<td>Netherlands (Kaatsheuvel)</td>
<td>219–20, 222</td>
</tr>
<tr>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Utrecht</td>
<td>Dutch</td>
<td>Netherlands (Utrecht)</td>
<td>107–8</td>
</tr>
<tr>
<td>Zeeland</td>
<td>Dutch</td>
<td>Netherlands (Zeeland province)</td>
<td>166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>------------------</td>
<td>----------------------------</td>
<td>-------------------------------------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>American English</td>
<td>USA</td>
<td>USA</td>
<td>22, 26, 34, 39, 70, 94, 114, 121, 123, 142–3, 147, 208, 238</td>
</tr>
<tr>
<td>British English</td>
<td>UK</td>
<td>UK</td>
<td>20–2, 26, 34, 39, 77, 95–6, 101, 122–3, 139, 145, 206, 218, 242</td>
</tr>
<tr>
<td>Canadian English</td>
<td>Canada</td>
<td>Canada</td>
<td>217, 222</td>
</tr>
<tr>
<td>Durham English</td>
<td>UK (Durham)</td>
<td>UK (Durham)</td>
<td>242</td>
</tr>
<tr>
<td>London English</td>
<td>UK (London)</td>
<td>UK (London)</td>
<td>235</td>
</tr>
<tr>
<td>Northern Irish English</td>
<td>UK (Northern Ireland)</td>
<td>UK (Northern Ireland)</td>
<td>118</td>
</tr>
<tr>
<td>NYC English</td>
<td>USA (New York City)</td>
<td>USA (New York City)</td>
<td>120</td>
</tr>
<tr>
<td>Old English</td>
<td>Extinct</td>
<td>Extinct</td>
<td>37, 90, 97, 142–3, 234</td>
</tr>
<tr>
<td>Scottish English</td>
<td>UK (Scotland)</td>
<td>UK (Scotland)</td>
<td>29, 71, 73, 114</td>
</tr>
<tr>
<td>Singapore English</td>
<td>Singapore</td>
<td>Singapore</td>
<td>99–100</td>
</tr>
<tr>
<td>Yorkshire English</td>
<td>UK (Yorkshire)</td>
<td>UK (Yorkshire)</td>
<td>242</td>
</tr>
<tr>
<td>Estonian</td>
<td>Uralic, Finnic</td>
<td>Estonia, Latvia</td>
<td>39</td>
</tr>
<tr>
<td>Etsako</td>
<td>Niger-Congo, Benue-Congo</td>
<td>Nigeria (Edo state)</td>
<td>155–6, 159</td>
</tr>
<tr>
<td>Ewe</td>
<td>Niger-Congo, Kwa</td>
<td>Ghana, Togo</td>
<td>71, 252</td>
</tr>
<tr>
<td>Fijian</td>
<td>Austronesian, Oceanic</td>
<td>Fiji</td>
<td>38</td>
</tr>
<tr>
<td>Finnish</td>
<td>Uralic, Finnic</td>
<td>Finland</td>
<td>7, 31–2, 39, 216, 247–9</td>
</tr>
<tr>
<td>French</td>
<td>Indo-European, Romance</td>
<td>Algeria, Belgium, Canada, France, Luxemburg, Monaco, Morocco, Switzerland; lingua franca</td>
<td>4, 12, 23, 26–7, 29–30, 37–8, 48, 58, 88, 90, 97–8, 116–17, 119, 124, 127, 133, 146, 178–9, 213, 219, 222, 253–4</td>
</tr>
<tr>
<td>Frisian</td>
<td>Indo-European, Germanic</td>
<td>Germany (Northeast); Netherlands (Friesland)</td>
<td>73–4, 135, 236–8</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Gaelic</td>
<td>Indo-European, Celtic</td>
<td>UK (Outer Hebrides)</td>
<td>133</td>
</tr>
<tr>
<td>Garawa</td>
<td>Australian, Garawan</td>
<td>Australia (Northern Territory; Queensland)</td>
<td>181–2, 193–4</td>
</tr>
<tr>
<td>German</td>
<td>Indo-European, Germanic</td>
<td>Austria, Belgium, Brazil, Germany, Italy, Liechtenstein, Luxembourg, Namibia, Poland, Romania, Ukraine, Russia, Slovakia, Switzerland</td>
<td>2–3, 7, 13, 23, 26–7, 29–32, 38–40, 42, 50–1, 71, 78–9, 81, 87, 101, 115–16, 127, 135, 149, 166, 169, 176, 185, 207, 209, 241–2, 249, 250</td>
</tr>
<tr>
<td>Central Franconian dialects (German)</td>
<td></td>
<td>Rhineland-Westfalia, Rhine-Palatinate; also Belgium, Luxembourg, the Netherlands (Limburg provinces)</td>
<td>169</td>
</tr>
<tr>
<td>Cologne German</td>
<td></td>
<td>Cologne</td>
<td>176</td>
</tr>
<tr>
<td>Giryama</td>
<td>Niger-Congo, Bantu</td>
<td>Kenya</td>
<td>172</td>
</tr>
<tr>
<td>Gothic</td>
<td>Indo-European, Germanic</td>
<td>Extinct</td>
<td>135</td>
</tr>
<tr>
<td>Greek</td>
<td>Indo-European, Hellenic</td>
<td>Cyprus, Greece</td>
<td>71, 78, 89, 209, 254</td>
</tr>
<tr>
<td>Haida</td>
<td>Isolate</td>
<td>Canada (British Columbia)</td>
<td>40–1</td>
</tr>
<tr>
<td>Hausa</td>
<td>Afro-Asiatic, Chadic</td>
<td>Burkina Faso, Niger, Nigeria</td>
<td>46, 213</td>
</tr>
<tr>
<td>Hawaiian</td>
<td>Austronesian, Oceanic</td>
<td>USA (Hawai‘i)</td>
<td>23, 31–2, 38–40, 46, 48–50, 53, 59, 136, 191</td>
</tr>
<tr>
<td>Hindi</td>
<td>Indo-European, Indo-Iranian</td>
<td>India, Nepal</td>
<td>22, 29, 233–5, 237, 239</td>
</tr>
<tr>
<td>Hixkaryana</td>
<td>Cariban</td>
<td>Brazil (Amazonas and Pará states)</td>
<td>46</td>
</tr>
<tr>
<td>Hopi</td>
<td>Uto-Aztecan</td>
<td>USA (Arizona)</td>
<td>136, 182, 190, 193, 196</td>
</tr>
<tr>
<td>Icelandic</td>
<td>Indo-European, Germanic</td>
<td>Iceland, USA</td>
<td>181–2, 189–90</td>
</tr>
<tr>
<td>Igbo</td>
<td>Niger-Congo, Benue-Congo</td>
<td>Nigeria (Southeast)</td>
<td>40, 95–6</td>
</tr>
<tr>
<td>Imonda</td>
<td>Border, Waris</td>
<td>Indonesia (Eastern Western Papua), Papua New Guinea (Sandaun province)</td>
<td>81</td>
</tr>
<tr>
<td>Indonesian (aka Bahasa Indonesia)</td>
<td>Austronesian, Malayic</td>
<td>Indonesia</td>
<td>48, 51, 59–60, 179</td>
</tr>
<tr>
<td>Inuit</td>
<td>Eskimo-Aleut</td>
<td>Canada, Greenland</td>
<td>2</td>
</tr>
<tr>
<td>Italian</td>
<td>Indo-European, Romance</td>
<td>Italy, USA</td>
<td>26–7, 31–2, 38, 79, 83, 137, 178–9, 185, 195, 201, 209, 211–12, 253</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------</td>
<td>--------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Itunyoso Trique</td>
<td>Otomanguean, Mixtecan</td>
<td>Mexico (Oaxaca state)</td>
<td>170</td>
</tr>
<tr>
<td>Japanese</td>
<td>Isolate</td>
<td>Brazil, Japan, Taiwan, USA</td>
<td>2–3, 5–6, 31–2, 38, 45, 52–3, 61, 101, 127, 160, 176–7, 179, 195, 253</td>
</tr>
<tr>
<td>Kalabari</td>
<td>Niger-Congo, Ijoid</td>
<td>Nigeria (Rivers state)</td>
<td>158</td>
</tr>
<tr>
<td>Karen</td>
<td>Sino-Tibetan, Tibeto-Burman</td>
<td>Myanmar</td>
<td>23</td>
</tr>
<tr>
<td>Kayardild</td>
<td>Australian, Non-Pama-Nyungan</td>
<td>Australia (South Wellesley Island)</td>
<td>77</td>
</tr>
<tr>
<td>Kikuyu</td>
<td>Niger-Congo, Bantu</td>
<td>Kenya</td>
<td>153–4</td>
</tr>
<tr>
<td>Kilivila</td>
<td>Austronesian, Oceanic</td>
<td>Papua New Guinea (Trobiand Islands)</td>
<td>76</td>
</tr>
<tr>
<td>Kipare (aka Asu)</td>
<td>Niger-Congo, Bantu</td>
<td>Tanzania (Kilimanjaro region)</td>
<td>168</td>
</tr>
<tr>
<td>Kisanga</td>
<td>Niger-Congo, Bantu</td>
<td>DR Congo (South)</td>
<td>157</td>
</tr>
<tr>
<td>Klamath</td>
<td>Penutian</td>
<td>USA (Oregon)</td>
<td>38–9, 237–8, 242</td>
</tr>
<tr>
<td>Konjo</td>
<td>Austronesian, Malayo-Polynesian</td>
<td>Indonesia (East Sulawesi)</td>
<td>48, 51, 59–60, 62</td>
</tr>
<tr>
<td>Korean</td>
<td>Isolate</td>
<td>North Korea, South Korea</td>
<td>26, 95, 97, 116</td>
</tr>
<tr>
<td>Koya</td>
<td>Dravidian</td>
<td>India (Telangana, Andhra Pradesh)</td>
<td>137</td>
</tr>
<tr>
<td>Kuki-Thaadow</td>
<td>Sino-Tibetan, Kuki-Chin</td>
<td>India (Northeast), Myanmar</td>
<td>164</td>
</tr>
<tr>
<td>Latin</td>
<td>Indo-European, Romance</td>
<td>Extinct</td>
<td>83, 253</td>
</tr>
<tr>
<td>Lenakel</td>
<td>Austronesian</td>
<td>Vanuatu</td>
<td>191</td>
</tr>
<tr>
<td>Limburgish</td>
<td>Indoeuropean, Germanic, Central Franconian</td>
<td>Belgium, Netherlands (Limburg provinces)</td>
<td>74, 169</td>
</tr>
<tr>
<td>Geleen Limburgish</td>
<td></td>
<td>Netherlands (Geleen)</td>
<td>74</td>
</tr>
<tr>
<td>Hasselt Limburgish</td>
<td></td>
<td>Belgium (Hasselt)</td>
<td>169</td>
</tr>
<tr>
<td>Luganda</td>
<td>Niger-Congo, Bantu</td>
<td>Uganda</td>
<td>84, 129–30, 135</td>
</tr>
<tr>
<td>Ma'ya</td>
<td>Austronesian, Malayo-Polynesian</td>
<td>Indonesia (West Papua, Raja Ampat archipelago)</td>
<td>176, 198</td>
</tr>
<tr>
<td>Macedonian</td>
<td>Indo-European, Slavic</td>
<td>Macedonia</td>
<td>197</td>
</tr>
<tr>
<td>Malay</td>
<td>Austronesian, Malayic</td>
<td>Indonesia, Malaysia</td>
<td>179</td>
</tr>
<tr>
<td>Ambonese Malay</td>
<td></td>
<td>Indonesia (Central Moluccas)</td>
<td>179</td>
</tr>
<tr>
<td>Malayalam</td>
<td>Dravidian</td>
<td>India (Kerala; Tamil Nadu)</td>
<td>129</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------</td>
<td>--------------------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Mandarin</td>
<td>Sino-Tibetan, Chinese</td>
<td>China (North); many other locations</td>
<td>38, 148, 162, 170</td>
</tr>
<tr>
<td>Henan Mandarin</td>
<td></td>
<td>China (Henan province)</td>
<td>38</td>
</tr>
<tr>
<td>Margi Mandarin</td>
<td></td>
<td>Nigeria (East)</td>
<td>76, 249</td>
</tr>
<tr>
<td>Mauritian Creole</td>
<td>Afro-Asiatic</td>
<td>Mauritius</td>
<td>58</td>
</tr>
<tr>
<td>Mba</td>
<td>French-based creole</td>
<td>DR Congo (Orientale province)</td>
<td>38</td>
</tr>
<tr>
<td>Mende</td>
<td>Niger-Congo, Mande</td>
<td>Sierra Leone</td>
<td>153, 233, 247</td>
</tr>
<tr>
<td>Mixtec Axcatlan</td>
<td>Otomanguean</td>
<td>Mexico (Southern)</td>
<td>163</td>
</tr>
<tr>
<td>Mixtec Mweria</td>
<td>Niger-Congo, Bantu</td>
<td>Tanzania (Southeast)</td>
<td>220</td>
</tr>
<tr>
<td>Nama</td>
<td>Khoisan</td>
<td>Namibia</td>
<td>34</td>
</tr>
<tr>
<td>Nambakaengo</td>
<td>East Papuan</td>
<td>Solomon Islands (Santa Cruz island, Temotu province)</td>
<td>76, 250</td>
</tr>
<tr>
<td>Ndunga</td>
<td>Niger-Congo, Ubangi</td>
<td>DR Congo (Equateur province)</td>
<td>38</td>
</tr>
<tr>
<td>Norwegian</td>
<td>Indo-European, Germanic</td>
<td>Norway</td>
<td>41, 43</td>
</tr>
<tr>
<td>Nubi</td>
<td>Arabic lexified creole</td>
<td>Kenya (Nairobi); Uganda (Bombo)</td>
<td>177, 198</td>
</tr>
<tr>
<td>Orochen</td>
<td>Altaic, Tungusic</td>
<td>China (Heilongjiang)</td>
<td>79</td>
</tr>
<tr>
<td>Oromo</td>
<td>Afro-Asiatic, Cushitic</td>
<td>Ethiopia</td>
<td>70</td>
</tr>
<tr>
<td>Pawaia</td>
<td>Trans-New Guinea, Teberen-Pawaian</td>
<td>Papua New Guinea (Chimbu province)</td>
<td>41</td>
</tr>
<tr>
<td>Persian</td>
<td>Indo-European, Indo-Iranian</td>
<td>Iran</td>
<td>176–7, 198</td>
</tr>
<tr>
<td>Pirahã</td>
<td>Muran</td>
<td>Brazil (Amazonas state)</td>
<td>38, 40</td>
</tr>
<tr>
<td>Plautditsch</td>
<td>Indo-European, Germanic</td>
<td>Russia (Western Siberia)</td>
<td>97</td>
</tr>
<tr>
<td>Polish</td>
<td>Indo-European, Slavic</td>
<td>Poland</td>
<td>39, 188–9, 250, 255</td>
</tr>
<tr>
<td>Portuguese</td>
<td>Indo-European, Romance</td>
<td>Brazil, Portugal, USA</td>
<td>27, 145</td>
</tr>
<tr>
<td>Brazilian Portuguese</td>
<td>Indo-European, Romance</td>
<td>Brazil, USA</td>
<td>145</td>
</tr>
<tr>
<td>European Portuguese</td>
<td></td>
<td>Portugal</td>
<td>27</td>
</tr>
<tr>
<td>Proto-Indo-European</td>
<td>reconstructed</td>
<td></td>
<td>148, 178</td>
</tr>
<tr>
<td>Provençal</td>
<td>Indo-European, Romance</td>
<td>France, Italy, Monaco</td>
<td>96</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------------------------------</td>
<td>--------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Rotokas</td>
<td>North Bougainville</td>
<td>Papua New Guinea (Bougainville province)</td>
<td>37–8, 40</td>
</tr>
<tr>
<td>Russian</td>
<td>Indo-European, Slavic</td>
<td>Russia, USA, former USSR</td>
<td>111, 243–4</td>
</tr>
<tr>
<td>Sanskrit</td>
<td>Indo-European, Indo-Aryan</td>
<td>India (Uttar Pradesh)</td>
<td>244</td>
</tr>
<tr>
<td>Sardinian</td>
<td>Indo-European, Romance</td>
<td>Italy (Sardinia)</td>
<td>224</td>
</tr>
<tr>
<td>Serbocroat</td>
<td>Indo-European, Slavic</td>
<td>Bosnia, Croatia, Serbia</td>
<td>255</td>
</tr>
<tr>
<td>Shiwilu</td>
<td>Kawapanan</td>
<td>Peru (Alto Amazonas province)</td>
<td>190, 196</td>
</tr>
<tr>
<td>Shona</td>
<td>Niger-Congo, Bantu</td>
<td>Mozambique, Zambia, Zimbabwe</td>
<td>158</td>
</tr>
<tr>
<td>Sinhalese</td>
<td>Indo-European, Indo-Iranian</td>
<td>Sri Lanka</td>
<td>29</td>
</tr>
<tr>
<td>Slovak</td>
<td>Indo-European, Slavic</td>
<td>Slovakia</td>
<td>223, 225–8</td>
</tr>
<tr>
<td>Spanish</td>
<td>Indo-European, Romance</td>
<td>Argentina, Belize, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Peru, Puerto Rico, Spain, Uruguay, USA, Venezuela</td>
<td>23, 29, 31, 38–9, 67, 97–8, 119, 127, 144, 175–6, 189–90, 222–4, 226, 228, 230, 235–6</td>
</tr>
<tr>
<td>Colombian Spanish</td>
<td></td>
<td>Colombia</td>
<td>38</td>
</tr>
<tr>
<td>Gran Spanish</td>
<td></td>
<td>Spain (Canary Islands)</td>
<td>222–4, 226, 228, 230</td>
</tr>
<tr>
<td>New World Spanish</td>
<td></td>
<td>Central and South America</td>
<td>235</td>
</tr>
<tr>
<td>Sundanese</td>
<td>Austronesian, Germanic</td>
<td>Indonesia (Java; Bali)</td>
<td>245</td>
</tr>
<tr>
<td>Swedish</td>
<td>Indo-European, Germanic</td>
<td>Finland, Sweden</td>
<td>96, 128, 175–6, 198</td>
</tr>
<tr>
<td>Tacana</td>
<td>Tacanan</td>
<td>Bolivia (La Paz province)</td>
<td>297</td>
</tr>
<tr>
<td>Tagalog</td>
<td>Austronesian, Malayo-Polynesian</td>
<td>Philippines</td>
<td>138, 201</td>
</tr>
<tr>
<td>Tahltan</td>
<td>Na-Dene</td>
<td>Canada (British Columbia)</td>
<td>256</td>
</tr>
<tr>
<td>Tamil</td>
<td>Dravidian</td>
<td>India (Tamil Nadu; Andhra Pradesh), Sri Lanka (North)</td>
<td>129, 137</td>
</tr>
<tr>
<td>Telugu</td>
<td>Dravidian</td>
<td>India (Andhra Pradesh)</td>
<td>84</td>
</tr>
<tr>
<td>Tera</td>
<td>Afro-Asiatic, Chadic</td>
<td>Nigeria (Gombe state)</td>
<td>138</td>
</tr>
<tr>
<td>Language</td>
<td>Family</td>
<td>Geographical area</td>
<td>Page numbers</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------</td>
<td>------------------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>Thai</td>
<td>Daic, Tai</td>
<td>Thailand</td>
<td>42, 46</td>
</tr>
<tr>
<td>Tharaka</td>
<td>Niger-Congo, Bantu</td>
<td>Kenya (Central)</td>
<td>153–4</td>
</tr>
<tr>
<td>Tiv</td>
<td>Niger-Congo, Bantoid</td>
<td>Cameroon, Nigeria</td>
<td>154</td>
</tr>
<tr>
<td>Tlingit</td>
<td>Na-Dene</td>
<td>Canada (British Columbia), USA (Alaska)</td>
<td>29</td>
</tr>
<tr>
<td>Toda</td>
<td>Dravidian</td>
<td>India (Tamil Nadu)</td>
<td>31</td>
</tr>
<tr>
<td>Tolitoli</td>
<td>Austronesian, Malayo-Polynesian</td>
<td>Indonesia (Sulawesi)</td>
<td>96, 104</td>
</tr>
<tr>
<td>Tonkawa</td>
<td>Isolate</td>
<td>USA (Oklahoma)</td>
<td>55</td>
</tr>
<tr>
<td>Tsou</td>
<td>Austronesian</td>
<td>Taiwan</td>
<td>23</td>
</tr>
<tr>
<td>Turkish</td>
<td>Altaic</td>
<td>Cyprus, Germany, Turkey</td>
<td>71, 82, 89, 93, 184–5, 209</td>
</tr>
<tr>
<td>Venda</td>
<td>Niger-Congo, Bantu</td>
<td>South Africa, Zimbabwe</td>
<td>71</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>Austro-Asiatic, Mon-Khmer</td>
<td>Vietnam</td>
<td>34, 46, 170</td>
</tr>
<tr>
<td>Welsh</td>
<td>Indo-European, Celtic</td>
<td>UK (Wales)</td>
<td>31</td>
</tr>
<tr>
<td>Weri</td>
<td>Trans-New Guinea, Kunimaipa</td>
<td>Papua New Guinea (Central province; Morobe province)</td>
<td>190</td>
</tr>
<tr>
<td>Wu</td>
<td>Sino-Tibetan, Chinese</td>
<td>China (East Central)</td>
<td>38, 170–1, 195</td>
</tr>
<tr>
<td>Shanghai Wu</td>
<td></td>
<td>Shanghai municipality, China</td>
<td>170–1</td>
</tr>
<tr>
<td>Yuhuan Wu</td>
<td></td>
<td>Zhejiang province, China</td>
<td>171</td>
</tr>
<tr>
<td>Xhosa</td>
<td>Niger-Congo, Bantu</td>
<td>Lesotho, South Africa, Swaziland</td>
<td>34</td>
</tr>
<tr>
<td>Yabem</td>
<td>Austronesian, North Huon Gulf</td>
<td>Papua New Guinea (Morobe province)</td>
<td>187</td>
</tr>
<tr>
<td>Yanyuwa</td>
<td>Pama-Nyungan</td>
<td>Australia, Northern Territory</td>
<td>82</td>
</tr>
<tr>
<td>Yoruba</td>
<td>Niger-Congo, Benua-Congo</td>
<td>Benin, Nigeria, Togo</td>
<td>46, 165–6</td>
</tr>
<tr>
<td>Zenzontepec Chatino</td>
<td>Otomanguean, Zatopecan</td>
<td>Mexico (Oaxaca province)</td>
<td>148</td>
</tr>
<tr>
<td>Zulu</td>
<td>Niger-Congo, Bantu</td>
<td>Malawi, South Africa, Swaziland</td>
<td>154, 160</td>
</tr>
<tr>
<td>Zwara (aka Zuwārah)</td>
<td>Afro-Asiatic, Berber</td>
<td>Libya (Northwest)</td>
<td>39</td>
</tr>
</tbody>
</table>
Subject index

ι [aɪəʊtə] see intonational phrase
μ [mjuː] see moras
φ [faɪ] see phonological phrase
σ [ˈsɪgmə]; see syllable
ω BrE [ˈoʊməɡə]; AmE [əʊˈmeɪɡə]; see phonological words

absolute slicing hypothesis 150
accents 159–61; non-stress vs stress 176–7;
pitch 163, 175
accentual analysis 160
adjustment, phonological 48, 63; adjustment
processes 49–54; adjustments in the native
vocabulary and 54; choosing between
rules and constraints in 58–62; constraints
approach 56–8; process of nativization
and 49–53; rules approach 54–5; starting
phonology and 48–9
affixes 8–9, 54, 192
affricates 30
airstream mechanisms 18
Aitken’s Law 71
alignment, tone 166–8
allomorph 98; morpheme alternant and
allophones 94
allophonic and stylistic variation 95–8
allophonic rules 101, 109–10
alveolar 29
alveolar ridge 25
alveolopalatal 95
ambiguity and non-specification, distinctive
features 72–5
ambsyllабic 142–4, 147
analysis, segment 43
Andersen, Torben 166
apical 77
approximant 30, 31; light or near-contact and
30
aspiration 20–2
assimilations 45, 89; coronal 233, 240, 256;
long-distance 243–6; nasal 240; place 239,
241; vowel harmony 247–9
association convention 151, 153–5, 172; accent
and 159–60
association lines 130
autosegments 130–2
back (tongue) 25
back vowels 26
Beckman, Mary E. 176
bilabial 29
bimoraic syllables 189
binary distinctive features 64, 66, 180
blade zone (tongue) 25
bleeding order 219
Booij, Geert 200, 209, 211
boundary: strength 200–1; symbols 85; tones
163, 201
boundedness and unboundedness in word
stress 182–3
bound morphemes 8
brace in SPE rules 88
bracketing, labeled 9
breathy voice 22, 23; tone and 170
catalexis 184
category membership 8
central vowels 26
Chatterji, Suniti K. 178
Chaucer, Geoffrey 103
Chen, Matthew 211
children: early language learning by 1–2;
learning sounds of their language in a
particular order 42; vowel production in 28
Chomsky, Noam 54, 76, 92, 111, 119
clashes 196–8
classes 9
Clements, George N. 25, 66, 128, 129, 133, 247
clicks 33–4
cliticization 203–4
clits 212–13
close-mid vowels 26
close vowels 26
coarticulation 201
coda 38, 126
coda-μ 189–90
cognitive factors 5
Cohn, Abigail 121
compensatory lengthening 134–6
complementary distribution 96
complex consonants 32–3
complex-place segments 249
complex segments 249–57; evidence for 251–2; representation of palatals and palatoalveolars 252–5
complex words 8
consonants 37; apical 77; cluster 38; complex 32–3; deletion, final 219, 222; double articulations 33; liaison 133, 143–4; manner-contour 33; non-pulmonic 33–4; number correlated to number of vowels 42; secondary articulations 32–3; segmental duration of 31–2; syllabic 39, 135
constraints: -based approach to adjustment 56–8; choosing between rules and 58–62; gradient violation and unranked constraints in 61–2; hierarchy 56; unranked 61–2
constricted glottis 70
constriction 28–31; major-class features 66–8; places of articulation and 29; types of 30–1 context in SPE rule 87
contour tones 151
contrastive underspecification 84, 242
contrastive vs redundant features 82–5
coronal 29, 76–7
coronal assimilation 233, 240, 256
counterbleeding: opacity 225; order 219–20
counterfeeding: opacity 223–4; order 219
creaky voice 22, 23; tone and 170
crown 29
decline, pitch 165
degenerate feet 188–92
delinking and spreading 238–41
dental 29
derivational affixes 9
derivational approach 54
devoicing 20–2; bleeding order and 219;
simultaneous-ordering theory and 92
diacritics 17
diphthongs 27, 44
dissimilation 83
distinctive features 11, 64, 75; ambiguity and non-specification 72–5; binary 64, 66; contrastive and redundant 82–3; feature values 66; laryngeal 66, 69–70; major-class 66–9; manner 66, 70–2; motivating 64–6; place 76–82; system 66
diversity, language 2–3
dominance in word stress 180–1, 188–92
Do-re-mi 3–4
dorsal 29, 78–82
dorsum see tongue
double articulations 33
downstep, tonal 164
dual structure 14–16; non isomorphic structures and 15
duplication problem 59
duration, segmental 31–2, 146
descent in word stress 180–1, 192–5;
extrametricality and 184
egressive speech sounds 18
jectives 34
Eliot, T.S. 128
Elsewhere Condition 216–17
English Pronouncing Dictionary 176
ergonomics, speech 44–5
exemplar models 122
extrametrical elements 184
extrinsic and intrinsic ordering 216–18
faithfulness constraints 56–7
feature geometry 232–3, 257–8; complex segments 249–57; implications of underspecification 241–9; long-distance assimilation 243–6; natural feature classes 233–4; spreading and delinking 238–41; tree diagram 234–8; vowel harmony 247–9
feature matrix 85
feeding order 218
final lengthening 201
final lowering of tones 165–6
flapping 143–4, 147, 217
floating 164, 166, 247
focus of change in SPE rule 87
foot 34; degenerate 188–92; dominance 180–1, 188–92; edges and iteration 192–5; and footing 179–83; independent evidence for 186–7; non-finality and clashes 196–8; shapes 188–92
formants 21
fo: perturbations 149; tone distribution and 170–2; see also fundamental frequency
free morphemes 8
fricatives 30, 44
friction 22, 30
front (tongue) 25
front vowels 26
fundamental frequency 21–2
front vowels 26
fundamental frequency 21–2

geminates 31
geminates and weight 195–6
generalizations: involving prosodic constituents 200–2; syllable-based 139–41
gestures 122; and gestural model 121–3
Gimson, A. C. 176
glotalic airstream mechanism 34
glottal stop 23; tone distribution and 171
open and vibrating glottis 18; closing 23; laryngeal features 69–70;
glottis 18; closing 23; laryngeal features 69–70;
harmony, vowel 247–9
Halle, Morris 54, 66, 224; on levels of representation 111, 119; on places of articulation 76, 92; on word stress 179, 185
Hammerstein, Oscar 3
hard palate 25
harmony, vowel 247–9
Harris, Zellig S. 111
hierarchical phrasal structure 9
hierarchy, constraint 56
high vowels 26
Hockett, Charles F. 111
Hooper, Joan B. 88
H raising 165
Hyman, Larry M. 154, 160, 174, 178
iambic structure 180–3, 188; edges and iteration 193–5; quantity-sensitivity bias 183; uneven 188
implementation, phonetic 109, 119–20; models of 120–2; of tone 162–8
implosives 34
inflectional affixes 9
information structure of sentences 204
initial strengthening 201
Inkelas, Sharon 208, 211
innate features of language 37
interpolations 120–1
intonational phrase 206–7
intonational tones 149, 166
inventory, lexical 113
Jacobs, Haike 253
Jakobson, Roman 41, 42
Jones, Daniel 176
Kahn, Daniel 88, 143
Kaisse, Ellen M. 213
Kaye, Jonathan 66, 224
Keating, Patricia A. 253
Kenstowicz, Michael 160, 185, 193
Keyser, Samuel J. 129, 133
Kiparsky, Paul 216, 222
labelled bracketing 9
labial-coronal segments 249
labial-dorsal segments 249, 252
labiodental 29
Lang, Walter 161
language(s): assimilation in 45; awareness of structure of 2; diversity 2–3; -external evidence 6–8; internal evidence 5; learning by children by age 5 1–2; system gaps 45–6; tone 23, 148–9, 178; varying complexity of 37–40
laryngeal features 66, 69–70
laryngeals and no manner or place feature specifications 73–4; laryngeals and no supralaryngeal specification 235–6
laryngealized voice 22
laryngeal node 235–6, 241, 244
larynx 17, 18–19
Lass, Roger 235
lateral plosion 135
lateralis 31
Leben, William R. 151, 158
Lehrer, Tom 4
level tones 151
lexical inventory 113
lexical and postlexical phonological rules 111–12, 124–5; application across word boundaries 114; controversial properties of 118–24; exceptions 112–13; native-speaker intuitions 114; non-derived environment blocking 118; reference to morphological tables 112; structure preservation 113–14
lexical and postlexical prosody 178–9
lexical representation 109, 248; surface representation and 118–24; underlying representation and 118–24
liaison consonants 133; ambisyllabicity and 143–4
Liberman, Mark 164, 179
linear order 91–3
linear representations 128; tones and inadequacy of 149–50
linguistic structures 1; awareness of 2; cognitive, phonetic, and social factors in 5; diversity of 2–3; morphosyntactic structure in 1, 8–10; observing 1–4; phonological structure in 1, 10–14; role of spelling in 4
linguists 4–8; language-external evidence used by 6–8; language-internal evidence used by 5
Lloyd, Joy 39–40
long-distance assimilation 243–6; vowel harmony 247–9
low vowels 26
lungs 17, 18–19; plosives and 30
Maddieson, Ian 42
major-class features 66–9
manner-contour consonants 33, 250
manner features 66, 70–2
marginal segments 43
match theory 212
maximal segments 222
maximal rule application 222
maximal surface transparency 222
maximum onset principle (MOP) 126–8, 209; syllabification rules, post- 141–6
McCarty, John 57, 74, 130, 192
melodies, word 149, 150–5
mid vowels 26
minimal pairs 95, 111
monophotongs 26–7
moras 3, 6, 136–9, 146–7; bimoraic syllables 189; moraic consonants 137; tonal analysis and 160
morphemes 8–9, 14–15; boundary 85; choosing the underlying form for 101–4; recurring elements in 11; segments in 11; tonal 157; variation in pronunciation of 94
morphological labels 112
morphology 8; effects and exceptions in word stress 185–6
morphosyntactic structure 1, 8–10; in every linguistic expression 14–15; morphemes and words 8–9; syntactic structure 9–10
motivating distinctive features 64–6
mouth 24, 25; nasals and 31; teeth 25, 29; uvula 25, 31; see also tongue
multivalued features 66
narrowing, vocal folds 22
narrow transcription 17
nasal cavity 24, 25, 236–7
nasals 30, 31
native-speaker intuitions 114
nativization process 49–53
natural feature classes 233–4
naturalness condition 65, 188–92
natural segment classes 64, 65
Nespor, Marina 200, 205, 211, 212
neutralization 95, 101, 109–10
Ngunga, Arminda S. A. 154
no crossing constraint 154–5, 172, 243
no crowding constraint 152–3, 172
non-derived environment blocking 118
non-finality and clashes in word stress 196–8
nonpulmonic consonants 33–4
obligatory contour principle (OCP) 132, 149, 158–9, 172
obstruents 30, 67, 70–1
Odden, David 158
onset 126
opacity 223; counterbleeding 225; counterfeeding 223–4; optimality theory and 223–30; rules and constraints 222–3
open-mid vowels 26
open vowels 26
operative level 50–1
optimal form 57
optimality theory (OT) 56–7, 159, 211, 215–16, 230–1; counterbleeding opacity as problematic for 225; counterfeeding opacity as problematic for 223–4; extrinsic and intrinsic ordering 216–18; feeding, counterfeeding, bleeding, counterbleeding 218–22; opacity and 223–30; stratal 225–30; stress and 187–98
organs of speech 17; lungs and larynx 17, 18–19
over-application 223
palatal 33; feature geometry representation of 252–5
palatoalveolar 29; distinguishing among 255–6; feature geometry representation of 252–5
parenthesis notation in SPE 89
peak 38, 39, 126
perceptual level 49–51
perturbations, f0 149; tone distribution and 170–2
pharyngeal 33
pharynx 24, 25
phonation 19, 22
phonemes 12–13, 94, 111
phonemic transcription 109
phonetic factors 5, 65–6
phonetic implementation 109, 119–20;
   deciding between phonology and 122–4; of
tone 162–8
phonetic stress 176; absence of 176–7; as
   obligatory feature 178
phonetic symbols 17
phonological context 94
phonological phrase 207–8
phonological structure 1, 10–14; as “built up”
   41; in every linguistic expression 14–15;
   extending beyond the syllable and foot 200;
similarities between different languages 36,
   47; writing rules 85–91; see also adjustment,
phonological
phonological utterance 205–6
phonological word 15, 209–10
phonology 36–7; deciding between phonetic
   implementation and 122–4; information
   in the lexicon 115–17; lexical 111–15;
postlexical 111–15; -sensitive syntax 201–2;
system gaps 45–6; universal 12–14, 37;
   universals and implicational relations 40–6;
   and varying complexity of languages 37–40
phrases, syntactic 9
Pierrehumbert, Janet B. 162, 164
Pike, Eunice V. 163
pitch 23
place node 234–5, 241, 245; secondary
   articulations 250–1
places of articulation 76; coronal 29, 76–7;
   dorsal 29, 78–82; features 76–82; labial 29,
   32–3; linear order 91–3; pharynx 29, 33;
radi cal 79
plosives 30, 44, 46
Postal, Paul 111, 188
postalveolar 29
postlexical phonological rules see lexical and
   postlexical phonological rules
postlexical prosody see lexical and postlexical
   prosody
post-MOP syllabification rules 141–6
preboundary lengthening 201
precompiled forms 213
prefixes 8
prenasalized stops 33
prepalatal 29
primary stress 34
Prince, Alan 57, 179, 192
privative features 66
privative tone 149, 168–9
pronunciation, variation in 94, 122; received
   pronunciation (RP) and 122–4
prosodic hierarchy 125, 200, 214; constituents
   205–10; deriving prosodic constituents
   211–13; factors determining prosodic
   phrasing 203–5; generalization involving
   200–2; intonational phrase in 206–7;
   phonological phrase in 207–8; phonological
   utterance in 205–6; phonological word in
   209–10; strict layer hypothesis (SLH) 202–3
prosodic phrasing factors 203–5
prosody, lexical vs postlexical 178–9
Pulleyblank, Douglas 154
pulmonic airstream mechanism 18
quantity (syllables) 136
quantity-sensitivity bias in word stress 183,
   188–92
radical 79
radical underspecification 242
ranking of constraints 56
received pronunciation (RP) 122–4
recurring elements in morphemes 11
recursiveness 9–10
redundant vs contrastive features 82–5
Reiss, Charles 224
representation, levels of 109; beyond surface
   118–24; defining an intermediate 109–11;
hierarchies and autosegments 128–33;
lexical phonology 109, 111–15, 124–5;
lexical representation 119; phonological
   information in the lexicon and 115–17
representations, underlying and surface 94–5,
   104–8; allophonic and stylistic variation
   95–8; arguments for two levels of 98–101;
   choosing the underlying form from 101–4;
   neutralization 95, 101
restructuring 204
resyllabification 141–6
retroflex 29
rhyme 38, 126
rhythm rule 207–8
Roca, Ignacio M. 242
Rodgers, Richard 3
root node 236
root, tongue 25
Rubach, Jerzy 116, 226
rule approach in adaptation 54–5; choosing
   between constraints and 58–62
rule interactions 218
Sagey, Elizabeth C. 66
scaling, tonal 163–6
secondary articulations 32–3, 250–1
secondary stress 34
segmental duration 31–2, 146
segments 11–12, 15; association lines 130;
 auto- 130–2; avoidance of complexity in
43; common versus unusual 40–1; complex
249–57; complex-place 249; distinctive
features (see distinctive features); duration
31–2, 146; in formation of syllables,
complexity of 38; manner-contour 250;
marginal 43; natural segment classes 64, 65;
phonetic symbols representing 17; skeletal
slots 129–30; writing rules 85
Selkirk, Elisabeth 147, 179, 200, 206, 211, 212
sentences 9; information structure of 204
Sezer, Engin 185
sign languages 12
Silverman, Daniel 49, 50
simplex words 8
simultaneous rule ordering 91–3
skeletal slots 129–30; unfilled and
unassociated 133
soft palate 25; nasals and 31
sonorants 30, 67
sonority profile 127–8
Sound Pattern of English, The (SPE) 54
speech: complex consonants 32–3;
constrictions in 28–31; devoicing and
aspiration 20–2; ergonomics 44–5; glottal
stop in 23; lungs and larynx in 17, 18–19;
nonpulmonic consonants 33–4; open
and vibrating glottis in 19; organs of
17; phonetic symbols and 17; pitch 23;
segmental duration 31–2; special types of
phonation 22; stress 34; vocal tract in 24–5;
vowels 25–8; waveform 20
spelling 4
SPE model 85–91, 215, 230; absolute slicing
hypothesis and 150; additional notations
88–91; brace notation 88; general
format 86–8; parenthesis notation 89;
transformational rule format 90; variable
feature values 89
spread glottis 69–70
spreading and delinking 238–41
stability, tone 149
Stampe, David 188
stem(s) 9
Steriade, Donca 83, 242
stops 30; glottal 23, 171; prenasalized 33;
voiced 222–3; voiceless 123
stratal optimality theory 225–30
stress see word stress
strict layer hypothesis (SLH) 202–3
structural change (SC) 87
structural description (SD) 87
structure preservation, lexical phonology
113–14
style 95
stylistic variants 95
subcategories 8–9
suffixes 8–9
supralaryngeal node 236–8, 241, 244
surface contrast 111
syllabic consonants 135
syllabification 126–8; autosegments 130–2;
compensatory lengthening 134–6; domain
127; hierarchies and autosegments 128–36;
liaison consonants 133–4; maximum onset
principle 126–8; post-MOP rules 141–6;
skeletal slots 129–30; sonority profile 127–8;
unfilled and unassociated slots 133–4
syllable(s) 11, 14–15, 126; ambi syllabicity
142–4, 147; -based generalizations 139–41;
bimoraic 189; coda 38; contact law 128;
degree of complexity in 38–9; footing 179–
83; moras 3, 6, 136–9, 146–7; onset 38, 39;
peak 38, 39; segments building 15; stressed
34; tone-bearing unit (TBU) 151–5, 159
syntactic residue 213
syntactic structure 9–10
syntax 8; phonology-sensitive 201–2
system gaps 45–6

- Tableaux in optimality theory 57
- Target-based phonetic implementation 122, 162–8
- Teeth: coronal articulation and 29; lower front
  25; upper front 25
- Timing tier 129
- Tonal analysis 160
- Tonal morphemes 1157
- Tone(s): accent and 159–61; alignment 166–8;
  association convention 151, 153–5, 172;
  -bearing unit (TBU) 151–5, 159, 164, 166;
  boundary 163, 201; contour 151; declination
  165; distribution and f0 perturbations
  170–2; downstep 164; final lowering 165–6;
  H-raising 165; inadequacy of linear model
  for 149–50; intonational 149, 166; languages
  23, 148–9, 178; level 151; obligatory
  contour principle (OCP) 132, 149, 158–9, 172;
  phonetic implementation of 162–8;
  privative principle 149, 168–9; scaling
  163–6; stability 149, 155–8; twin
  sister convention 156, 172; undershoot 166;
  upstep 163; voice quality and 170; word
  melodies and 149, 150–5
tongue 25; body features 78; see also dorsal articulation; mouth
transcription, narrow 17
transformational rule format 90
transparency and opacity: rules and constraints 222–3
tree diagram and feature geometry 234–8
trills 31
trisyllabic laxing 113, 118
trochaic structure 180–3, 188; edges and iteration 193–5
Twaddell, W. Freeman 111
twin sister convention 156, 172

unary features 66
under-application 223
undershoot, tonal 166
underspecification: implications 241–9; theory 83
undominated constraints 59
uneven iamb 188
unfilled and unassociated slots 133
univalent features 66, 76
universal phonology 12–14, 37
universals 40–6
unmarked pronunciation 56
unranked constraints 61–2
upstep, tonal 163
utterance, phonological 205–6
uvula 25; trills and 31

vacuous rule application 87
values, feature 66
variable feature values 89
velaric airstream mechanism 33
velarization 33
velum 25
Venneman, Theo 88, 127, 128
Vergnaud, Roger J. 179
vocal cords see vocal folds
vocal folds 18; devoicing and aspiration 20–2; open and vibrating glottis 19; in producing vowels 25; special types of phonation 22; tone distribution and 170–2
vocal tract 24–5

Vogel, Irene 200, 205, 211, 212
voice 19; breathy 22, 23, 170; creaky 22, 23, 170; devoicing and aspiration 20–2; laryngeal features 69; laryngealized 22; onset time 20; pitch 23; quality and tone 170
voiced stops 222–3
voiced vowels and consonants 19
voiceless sounds 19, 23; fricatives 44, 71; plosives 44, 46
voicing: special types of phonation 22; whisper 22, 23
vowel(s) 25–8, 37; common versus less common 45; diphthong 27, 44; harmony 247–9; monophthong 26–7; nasalization 27–8; number correlated to number of consonants 42; production in children versus adults 28

weight-by-position (Coda-µ) 189–92, 193
weight in word stress 180–1; geminates and 195–6
weight-to-stress principle (WSP) 189–92, 193
Weijer, Jeroen van de 253
whisper 22, 23
Wistrand, Kent 163

word(s): boundary 85, 114; feet 34; melodies 149, 150–5; morphemes and 8–9; phonological 15, 209–10
word stress 34, 174–5, 198–9; absence of phonetic stress 176–7; degenerate feet and 188–92; describing patterns of 179–87; dominance in 180–1, 188–92; edge in 180–1, 184, 192–5; flexible right edges 184; footing 179–83; foot shapes and 188–92; independent evidence for the foot in 186–7; lexical vs postlexical prosody 178–9; manifestations of 175–9; morphological effects and exceptions 185–6; non-finality and clashes in 196–8; optimality theory and 187–98; phonetic stress as obligatory feature 178; phonological features of 174; weight and 180–1, 195–6
writing rules 85–91; boundary symbols 85; feature matrix 85; spreading and delinking 239–41