Phonology

Marc van Oostendorp

July 31, 2024

Contents

The subject matter of phonology

1.1 Where in the world is phonology?

When a character in a comic story gets very angry, she may start shouting something which appears in a little cloud above her head and reads as '💣⚡⚠☠☹🦇!' Those symbols undoubtedly satisfy the communicative needs of characters in comics rather well: every reader understands immediately what intention they are supposed to reveal. We also can easily imagine some sound — a deep grunt or growl — which is represented by such a list of characters and which would satisfy the same communicative goal.

It would be possible to devise a communication system which consists entirely of sounds every human being can make, and there are thousands of those, but which are difficult to notate precisely in an alphabetic writing system, even if we extend it by a long list of emoji. Human beings have a sufficiently large repertoire of sounds: we can whistle, hiss, snort, clap our hands, and do thousands of other things to produce noise with our bodies. Each of those sounds we could call a word, and then combine these words into sentences in the way in which we know syntax of human language does such things. Words would be literally atoms of sound and meaning.

Such a language would be convenient for a number of purposes. All words would be easily distinguishable: the chances that two words would sound as similar as *pin* and *pen* in English would be very small. So we could minimize confusion. At the same time, the words would not be too difficult to make: a human body would be enough, and most people happen to always have one at their disposal.

Yet, as far as we know, no human language works in this way. Rather, the words in every language that has ever been described seem to have much more fine-grained internal structure. For instance, all languages seem to dispose of a finite list of vowels and consonants, and all words consist of combinations of those — the smallest number of known inventories of consonants and vowels is about 11, the largest about 150.

Obviously, when phonologists talk about vowels and consonants, they talk about sounds, not about letters, which are typically seen as mere graphical representations of those sounds. Phonologically speaking, English does not have a consonant *c*. The letter of that shape, <c>, is sometimes pronounced as [k] and sometimes as [s].

You will have noticed that I put orthographic letters in angled brackets and phonetic symbols in square brackets. This is common practice within the sister disciplines of phonology and phonetics. Since, in spite of all technological advances, it is still impractical to incorporate real sounds in a book such as this one, we need to do with symbols representing them. For this I use the so-called IPA system, a 'phonetic alphabet' consisting of many more letters than the normal Latin alphabet and which is explained in section 1.2. IPA is a real standard; it is presently used by virtually every linguist in the world who describes sound systems.

It is very important to remember that phonology is about sounds and not about letters. Many students confuse these two in the beginning. In an alphabetic system, letters are a way to write sounds, but we study only the sounds. In the history of mankind, writing is a fairly recent invention, and still a majority of the languages spoken today has never been written, and for those which are written there may still be a large number of speakers who are illiterate. Nevertheless, in some case, we do need to talk about orthographic letters; to distinguish them from IPA symbols, we use the angled vs. square brackets.

The subject matter of phonology

phonology

The discipline which studies the sound systems of language is called *phonology*, in which we can recognize the Greek words φονή [fonɛ] which means 'sound', and λόγος [loɣos], which means 'study'.

The description of the phonology of a language usually starts with making an inventory of all the basic elements. These are the consonants and vowels of the language, for instance, but there can also be other elements, which may not always be transcribed in the orthographic system. For instance, the following two sentences sound different in English:

(1) a. You have eaten already.
    b. You have eaten already?

The statement in (1a) sounds different from the question in (1b) because the two sentences are pronounced with different so-called pitch contours: questions typically end in a higher pitch. Also these tonal differences count as part of the set of primitive elements of the language. (As we will see in chapter 4, in many languages not just sentences, but even individual words can be distinguished by the pitch at which they are pronounced.)

If we would restrict ourselves to just setting up inventories of vowels and consonantss, phonology would probably not be a very exciting discipline — it would mostly involve a very low level bookkeeping activity. Making such inventories however is just a very first and basic step in understanding how sound systems are organized in human language. Fortunately, there are more challenging and interesting questions which immediately rear their heads.

In the first place, it turns out that the inventories of sounds are not just unstructured lists of consonants, vowels, tones, and possibly some other elements. They have much more internal structure in that certain kinds of sounds seem to group together in certain ways, and furthermore such structures are very similar from one language to the next. For instance, if a language has the three consonant sounds [m,p,b], they often show a similar kind of behaviour, for instance they appear in similar positions in the word. This is because they are all *labial* sounds, formed with the lips. Apparently labiality is a factor that can play a role in the organisation of sound systems.

Once we observe this, we have two research questions: what *is* this common structure of sound inventories in languages of the world? And what would explain that those commonalities exist? A lot of debate in the scholarly phonological literature is about these issues, and we will return to several aspects of that debate in the following Chapters.

Another type of question arises because sounds tend not to be stable, but are realized differently in different consonants. A consonant or vowel, or even a sequence of consonants and vowels, might have a slightly different shape in one context than in another. Take, for instance, the English plural suffix. This has three different phonological shapes:

(2) a. -[ɪz]: *fishes, passes*
  b. -[s]: *cats, parks, lamps*
  c. -[z]: *lambs, pans, beds*

When I write sounds between square brackets, I refer to the *International Phonetic Alphabet*. This is a convenient way in which all sounds of all languages can be represented. I assume in this book that you know about the International Phonetic Alphabet, but if you don't you can easily look it up on the internet.

International Phonetic Alphabet

These shapes are dependent on the phonological make-up of the stem. If the latter ends in a [s], [z] or [ʃ] (represented by <sh>), it has the shape [ɪz] <es> (2a). If the stem ends in a so-called voiceless consonant like *p, t, k*, the suffix is the equally voiceless *s* (2b). Otherwise, it assumes the voiced form [z] (2c). (We will return to this in section 3.1).

Again, these alternations show many similarities with phenomena we find in other languages of the world (for instance many of them refer to the difference between voiced and voiceless sounds, and in many of them a [ə] or [ɪ] sound play a special role), so that again two questions arise. First, which kinds of alternations are available in languages of the world, in other words, which sounds can alternate with which other sounds? And secondly, what explains the fact that these alternations exist and that they seem to come from a rather small repertoire of possibilities? What does this tell us about human cognition, how people influence each other, how they organize their main system of communication? Debates about such questions are the bread and butter of phonological research, and after studying this textbook you should be able to follow such discussions.

Phonology and phonetics

There are many more things to be heard in a human speech utterance. For instance, (1a) would sound different if it were pronounced lovingly than in anger, different when pronounced by a child than by an adult, different when produced in a noisy environment rather than in a silent one. Such differences between moods or physical bodies are undoubtedly interesting and they have led to a lot of insightful research, but they are *not* studied in phonology.

Phonology is the study of sounds as part of language as a formal system for expressing and communicating thought, and that is where we typically find the answer to phonological questions: we only count those sound differences which are linguistically relevant, for instance because they correspond to a difference in linguistic meaning between two words, a difference in semantics. In that sense the sound difference beteen *e* and *a* is linguistically relevant in English, because *bet* and *bat* are different words with a different meaning. The difference between questions and statements is also linguistic, but the difference between loving and hating the person you speak to is not, at least not in the same way.

When doing science, we divide the world up in small subdomains, which we believe we can study independently from each other. This does not necessarily mean that reality can be cleanly divided up like this in everyday life, but it has proven to be the only way to get to some scientific understanding of the world if we sometimes abstract away from certain aspects. Sounds are extremely complex, humans are extremely complex, we may only get hopelessly confused if we try to understand everything at the same time. That is why we need fields like linguistics to say: we concentrate only on those aspects of reality that we call *language*; and subfields like phonology to say: we concentrate only the smallest elements of language, defined in this way. Phonology is distinguished from several neighbouring disciplines for this reason — which still overlap in their empirical domain, so that most phonologists will usually have a working knowledge of at least some of the other disciplines.

phonetics

One of these is *phonetics*. The differences between phonology and phonetics are subtle and complex — some scholars even suggest that there is no distinction between them at all. But usually it is assumed that phoneticians study the physical and physiological aspects of speech sounds: they use technology to study such sounds as part of the natural world. Phonology on the other hand studies sounds as part of language as a system, and the cognitive or social structures which underlie this system. In terms of a classical philosophical dichotomy, phonetics is about the body and phonology is about the mind; phonetics is a natural science, and phonology a cognitive science.

The issue is obviously very complicated, if only because it has proven to be notoriously difficult to disentangle mind and body. If I say the word *pen*, at some point my mind has to translate this word into a set of instructions to my lips and tongue and other muscles — what is the moment in which we make a transition from my intention to the physiology, from the mind to the body? Fortunately, in the everyday business of linguistic analysis we can take a pragmatic attitude towards this question: everybody uses the tools they have to study what they can. In any case, it is not really possible to study phonology without some basic knowledge of phonetics. I will therefore provide some of this knowledge in the chapters to come, although I encourage you to also learn more about phonetics, which is an extremely interesting discipline providing lots of fascinating insights. (At the end of every Chapter I provide a list of suggested literature.)

gradient

categorical

There are also differences between typical phonetic and typical phonological *data*. Since phonetics deals with the messiness of the physical world, a lot of its data are *gradient* (they involve values which can be written as real numbers, like 1.34253... or 65.696969...: no matter how precise we are, we can always imagine a little bit more precision), whereas phonology, which deals with the more abstract and therefore cleaner system of the human mind is assumed to be *categorical*. A consonant in a language is a *p* or an *b*, but not 13.4% a *p* and 86.6% a *b*. When phonology deals with numbers at all, they are therefore usually (discrete) *natural* numbers: $0, 1, 2, ...$; but we will see that numbers hardly play any role in most phonological analysis at all.

Another difference is that phonetics is not exclusively a linguistic discipline. This means that on the one hand, phoneticians study certain topics which phonologists would not consider. The differences between loving and angry speech or between younger and older speakers, of of speaking in different places, mentioned above, are cases in point. Generally, the speech signal contains much more pieces of information than purely linguistic ones. From the way somebody speaks, we can draw conclusions about their body size, age, emotional state and gender. We can hear how tired or agitated the speaker is, and according to some engineers one can even hear whether somebody speaks the truth or not. (Those engineers are involved in building polygraphs.) Such topics are typically studied by phoneticians, and not by phonologists.

I already mentioned that some people deny that there is a distinction between phonology and phonetics. The most common form this point of view takes is assuming that also cognitively people deal with phonetic representations, and that is all there is. Even if this would turn out to be true, the generalisations discussed in this book still would be true and in need of explanation.

Phonology and morphosyntax

Phonetics is obviously further removed from core linguistic disciplines, such as *morphology* and *syntax* than phonology is. Morphology is the study of the way in which words are built out of other words — how *books, bookish* and *bookcase* are all formed on the basis of *book*. Syntax is the study of how words are combined into sentences. (Because the differences between the two is often immaterial from a phonological point of we, we sometimes speak of *morphosyntax*). Both of these are grammatical disciplines, and so is phonology: they all describe the knowledge of the speakers of a language and deal with the systematicity in individual languages, or in 'human language' as a category.

morphology
syntax

It is said that one key property of human language is that it has *double articulation* (also sometimes called *duality of patterning*): every language consists of an inventory of small elements without meaning, such as consonants and vowels. These can be combined together to form words and affixes, which are (mostly arbitrarily) associated to a meaning. That is the first articulation or patterning. These smallest meaningful units can then be combined into larger units, like more complex words or sentences, in which also the organisation (the syntax) adds to the meaning. That is the second articulation.

double articulation

This description suggests that the first articulation is hierarchically below the second one (it stops below the level of the morpheme), but that is not correct. It seems better to view the two patterns in parallel (this is further discussed in Chapter 8).

The sounds of language are also sensitive to the higher levels of morphosyntactic organization. In the phonological alternations we mentioned above — of which the different shapes of the English plural suffix were examples — this already became apparent: phonology should be able to somehow 'see' the structure that is built by the morphology. In other languages, it can also see the syntactic structure in the same way.

Every linguistic utterance thus has the two patterns at the same time: on the one hand, it consists of meaningless sound symbols, that are somehow strung together to produce the sound stream we hear when somebody speaks, but at the same time — as it were in a different dimension —, it organizes the meaningful unit in a parallel way.

allomorphy

Just like with phonetics, the boundaries between phonology and morphosyntax are not always clear. One issue is *allomorphy*: morphemes can have a different shape depending on their context, and we can argue whether the resulting pattern should be studied in morphology or in phonology. Here is an example from Kalkatungu (an extinct Pama-Nyungan language from Australia). The genitive in this language is expressed by *-ku* if the stem ends in a consonant, and by *ja* if it ends in a vowel:

(3) a. *t̪uat-ku* 'snake', *upun-ku* 'frog', *t̪untal-ku* 'moon'
b. *macumpa-ja* 'moon', *ntia-ja* 'snake', *kupu-ja* 'spider'

We have seen another example of allomorphy in (2). There, we presented the different shapes of the English plural suffix as evidence for phonology: the [z] sound of the suffix changed its shape based on the phonological environment. That is the reason to consider this pattern to belong at least partly to the domain of phonology. There does not seem to be a reason to do the same for the Kalkatungu case: the shapes *ku* and *ja* sound too differently from each other to be related in a sensible way, and there also is no clear phonological reason why one would be chosen after a consonant and the other after a vowel.

For this reason, most linguists think that the Kalkatungu alternation is not part of phonology, but of morphology. But quite obviously, the boundaries are not necessarily always clear: two allomorphs might look *somewhat* similar and there might be a reason why they have this shape, but that purported reason might be a little bit far-fetched. Notice that the Kalkatungu alternation still refers to a phonological property of the stem, viz. whether it ends in a vowel or a consonant.

Again, most phonologists take a pragmatic approach to these questions. We deal with those phenomena that we can account for in our theories and leave other data to related fields which might be better equipped to deal with them. The English data can get a plausible explanation in the phonology, so we consider them as phonological; the Kalkatungu data on the other hand have to be solved by morphology. The problem arises much less with respect to syntax; there are very few phenomena where it is unclear whether they are syntactic or phonological; more discussion on this will be provided in Chapter **??**.

Where in the world is phonology

If phonologists study the structure of sound systems in human language, they must obviously assume that there is such a system to be found. The question then arises what is the reason that languages have structured sound systems to begin with.

We can only answer this question if we have an idea about the ontology of language: where in the world do we locate this phenomenon that is the object of study for linguistics? There are roughly three possible answers (each of them can be refined in a number of ways).

The first is that language is an abstract entity, which we can study independently of its speakers, much like the way many people study mathematics independent of how real humans calculate. We can call this the *Platonic view*, after the Greek philosopher Plato. Many linguists have taken this view in the past — it is for instance the basis of well-known metaphors of language as a living organism or of Portuguese being a daughter language of Latin. Under this view, the fact that language is logically structured is inherent in the fact that it exists: for a Platonist, abstract objects will always have a structure that is worth studying. **Platonic view**

The second point of view sees language as a cognitive object, as something that is ultimately represented in the brain or the mind of an individual. We can call this the *psychological* or *cognitive view* of language. It implies that we ultimately believe that the structure of language is causally related to the structure of human cognition. The way a phonological inventory is organized can under such a view for instance be understood in light of the way memory works in the human brain. Inversely, the study of phonology can then tell us something about the internal structure of the machine we call the mind. **cognitive view**

This second view has been the dominant one in phonology (as well as in theoretical syntax and morphology) for quite a long time. It was championed by the famous generative linguist and philosopher Noam Chomsky (1928), but also many linguists who take very different views on the structure of language, on linguistic methodology, etc., implicitly or explicitly see language as an essentialy cognitive object. Since this is the dominant view, it also is the one that is in the background of things I write in this book.

The third view sees language primarily as a social object, as a phenomenon that does not belong to one individual, but always to a group of people, for instance because it emerges if people communicate. We can call this the *sociological view* of language. It implies that the structure of language is caused by the way in which human interaction works. Although this way of thinking has not been as dominant as the cognitive view, it has always been there, and if I am not mistaken, it is gaining ground. A lot of sociolinguistic work obviously takes this point of view, as does a lot of historical linguistics, but in recent years there has also been interest in building computer models in which it is shown that models with a small number of computer 'agents' communicating with random sounds in the course of time start converging on sound systems which look like human language. **sociological view**

We cannot decide here which of these points of view is the 'correct' one. Probably they all are, since they are not mutually exclusive. Language is somehow encoded in the human brain *and* it is used in human interaction. Apart from this it may also be an abstract object with mathematical properties. In that case we will have to figure out which aspects of phonology are best explained from which basic principles. It *is* however very important to recognize that there are these different points of view and that what an individual analyst sees as a convincing argument may depend on his or her point of view.

There is yet another way to see the place of phonology in the world. This is by considering the distinction between *synchronic* and *diachronic* phonology — a distinction which is due to the Swiss linguist Ferdinand de Saussure (1857- **synchronic** **diachronic**

1913). We can see language as a static object, which exists at a given point in time (15th Century English, Swahili as it is nowadays spoken in Kenya): in that case we are doing synchronic linguistics. We can also study the dynamics of language as it is (continuously) changing; that is the subject matter of diachronic linguistics.

Again, it does not make sense to say that either of these views of language is the 'correct' one. It is an undeniable fact that languages (at least if seen from a sociological view) are subject to change. But the same is of course also true for objects in the physical world, which are constantly moving around, decaying, etc., which does not mean that we cannot study the structure of a physical object such as, say, the human hand as if it is stable — as a matter of fact something which is really constantly moving cannot be studied at all.

Most of phonological theory nowadays is synchronic: explanations for patterns are sought in structures which are treated as stable. This is definitely true for the style of theorizing which is presented here. However, there are also many phonologists who believe that many, most or all patterns can, and even should get a diachronic explanation. Again, it does not make sense to stipulate that one point of view is correct and the other one is wrong, but arguments and data will be seen in a different light from different points of view. Images of the brain when somebody pronounces a word are not necessarily interesting for somebody interested in diachrony; historical data about the way in which a word was pronounced in the past do not always contribute to our understanding of synchronic linguistics. This leads us to the point of which methodology phonologists nowadays should follow.

1.2 Phonological data and methodology

Since it is concerned with finding patterns, the study of phonology, like every scientific enterprise, requires a level of abstraction.

This emphatically does *not* mean that the phonologist does not deal with empirical data. Quite to the contrary, there are many sources of information that a modern phonologist can use to discover the abstract patterns. It is therefore important to understand these different types of data, and to be able to use different scientific methodologies to discover them.

In this section, we discuss the five most important types of data, and data collection methods, used in comtemporary phonology. It is important to realize that each of them is important and shows a piece of the puzzle; but that they also all can have problems.

There is the important anecdote of the people who encounter an elephant in the absolute dark, and each starts feeling with their hands. The one who feels the trunk thinks they are dealing with a snake, the one that feels one of the legs thinks this is a tree, etc. Scientists are similarly dealing with an unknown object, and it is absurd to believe that there is one given way to discover what it is.

Introspection and fieldwork

Classically, phonologists (and linguists more generally) have collected their data about a language by asking native speakers whether something was said like this in their language or not (their 'judgement'). Sometimes the researcher would be their own informant, when they were working on their own language. But in many cases, they also worked with an informant who was not necessarily a linguist to discover the sound system of a given language. If one works with one's own judgements, this is called *introspection*; if one works with other people's judgements, this can be called *fieldwork*. The advantage of fieldwork over one's own judgement is that the judgements are not influenced by the speaker's knowledge of what the theory predicts and therefore more objective; the advantage of introspection is that of course nobody is as patient and as willing to ponder new examples as long as the researcher.

introspection
fieldwork

How to conduct fieldwork? One simple thing one can obviously do, is try to record as many words of the language as one can find. One problem then is how to write out the sounds of the language. First, we need some kind of writing system. The orthographies of existing languages, say, English, do not suffice, because they do not faithfully record every sound distinction of the language itself, let alone of more foreign languages.

For this reason, linguists use the so-called *International Phonetic Alphabet* (*IPA*), a convention for writing all the sounds of all (known) languages in the world, which I have already briefly introduced above. The IPA consists of a large number of symbols and diacritics which together allow the researcher to go into as much detail as is useful for linguistic analysis. It has been tested on many languages and only very occassionally has it been extended because a newly discovered language seemed to possess a sound which was nowhere else to be found. It is very important for phonologists as well as for phoneticians to be able to read and write transcriptions in IPA. Here is an example of a sentence transcribed in IPA:

International Phonetic Alphabet
IPA

(4) ðɪs ɪz ən əgzæmpl̩ əv ə sɛntəns trænskraɪbd ɪn aɪpiːeɪ

Alternatively, one can of course decide not to write out any of the sounds at all, but make audio recordings. This may definitely be preferable while being in the field, if only because of course you will always miss something while transcribing. However, it is common practice to present IPA transcriptions in analyses and publications, so far only rarely accompanied by sound files. Furthermore, it is the case that every language apparently *can* be transcribed in the IPA, and finding which are the vowels and the consonants is one of the phonologists' tasks. So making a transcription is basically the first step towards an analysis.

Transcribing random words is obviously not the most ideal way of doing research if we are trying to find patterns. It is possible, for instance, that some sounds of combinations of sounds occur only very rarely, and we happen to miss them in our sample. Various methods have been developed to find the patterns we are looking for most systematically. Many of these will be studied in the next chapters, in which a lot of data are based on fieldwork.

The data we study as phonologists are of course not always from our own fieldwork or introspection. Science is a cumulative enterprise: we build on the work of other people, who have in turn built on the work of yet other people. A lot of data have been collected in this way by other researchers over the course of the past decades, and these play an important role in the literature since then. This does not mean that you should take such data without criticism, because it is always possible that mistakes have been made. The trick is to find the balance between always being critical of our sources and knowing when to trust what other people have said. Since it is impossible to do intensive fieldwork on more than a handful of languages over the span of a lifetime, we have to trust other people's work if we want to draw conclusions which go beyond such a small number of data points.

It should be noted that fieldwork is often done with representative adults, but phonologists have also worked with specific groups, for instance children who are acquiring their language, aphasic patients who have lost part of their language, second language learners, etc. However, in many of such cases experimental work (see below) is more common.

Corpora and databases

corpora

spontaneous speech

Another means of discovering patterns is by studying large data sets, for instance in the form of *corpora* — large, typically electronic, collections of natural language data. These can be for instance transcriptions of conversations or of monologues, what is usually called *spontaneous speech*, but also of more structured interviews based on a questionnaire.

Such corpora potentially give insight in different dimensions of natural language. For instance, they can tell us something about how frequently words, or syntactic constructions, or individual sounds are really used in everyday language. Also, the way in which people really speak might be different from the way in which they *think* they speak — the latter is the kind of data we obtain with the methodology described in the previous section. Finally, corpora are also more easily verifiable than judgements — a scholar can put the data they have used online, so that other people can verify them, or even use them for other purposes.

A special type of corpus worth mentioning is the sociolinguistically annotated corpus, which provides information about the background of the speakers; typically their age and gender but also other information about their position in society (for instance, where they were born, their occupation, the way they identify themselves, their lifestyle). These are important in particular if we take a sociological point of view of language, as described above and consider a language to be a property of a human social group.

We know that any human social group of some size consists of many subgroups: men speak slightly differently than women, etc. Given this fact, many sociolinguists propose that the language use of an individual speaker can only be understood as a function of the language system of the various subgroups they belong to. Under such assumptions, studying a corpus without the relevant corpus information does not make a lot of sense.

On the other hand, working with corpora also has several drawbacks. One of them is that it is very difficult to build a corpus that is really balanced and, above all, representative. When we record people, they will almost always be influenced by that fact and speak differently than when they speak in their ordinary voice.

A more important problem of using corpora is that certain patterns might just not occur even though they are theoretically possible — they are just exceedingly rare. But it might be these very rare patterns which show an important piece of the puzzle. As a matter of fact, it is sometimes argued that such patterns give a good indication for what is 'really' going on. If people can distinguish between 'good' and 'bad' patterns which are equally rare, they cannot have made this judgement based purely on the data they have encountered. There must be some other factor at work — for instance, innate knowledge in the sense of Chomsky (1968, 1986, 1995). But also other 'external' sources of language patterns, for instance generalization capacities might be justified by such evidence.

Also, although corpora give us insight in how people speak, it is not clear that they show us how people *intend* to speak, and one might argue that the object of study for phonology is the latter. People make slips of the tongue, are confused, do not speak very clearly, say things which for some reason do not really fit into their language system. If it is the language system we are after, it is difficult to find out what people intended; and for this we have to ask them, i.e. use (also) the fieldwork methodology.

A final problem with many linguistic corpora which is not inherent to the technology is that linguistic databases often consist of transcriptions only, and those often even in an orthographic form. It is obviously sometimes difficult to deduce from these how the data *sound*, let alone what they tell us about the phonological organization of the language in question. There are, however, also corpora which include phonological or phonetic transcriptions; and some of them even have sound files. (We will discuss examples in later chapters.)

Another type of electronic data source are *databases*; the difference with corpora is that databases give more structure to the data. A corpus consists of (transcriptions) of texts. Extra information about individual text items such as words might be added, as well as information about the text as a whole. Databases consist of more structured information, for instance about whole languages: lists of sounds or sound combinations, words, morphological paradigms, etc. Also, while corpora tend to concentrate on a single language, many databases are concerned with typology, i.e. comparing languages in a certain dimension. A well-known example of a phonological (or phonetic) database is the UCLA Phonological Segment Inventory Database, which contains tables of all the vowels and consonants in 451 languages — it is usually estimated that at present there are about 6,000 languages spoken in the world. Using this database, we can investigate claims such as 'all languages have a consonant [t]'. **databases**

An advantage of databases is that they are relatively easy to search for such claims. Furthermore, if we have a lot of data, we can apply statistical techniques which can filter out individual errors. A problematic aspect of databases is that they will often be eventually based on the fieldwork methodology described in the previous subsection. This means that there may be many errors, differences in interpretation of the data, etc. inherited from that fieldwork, and since databases are in such cases based on the work of many different, often independent, fieldworkers, it is very difficult to filter out all those errors.

Artificial evidence

Sometimes the kind of data we find in the world of everyday speech is not enough. In such cases, linguists may also use *artificial* data, which has been consciously created for some reason.

There are many types of such artificial data. Some of these have been created by non-linguists. One example of this is poetry; poets traditionally play with the sound structure of language by using rhyme, alliteration and other means. Especially in the case of historical data, such sound patterns are often the only indication about the phonology we have. For instance, from the fact that two words are put in rhyming position, we may conclude that they ended in the same or similar sounds, even though they were written differently. Similarly, what we know about the difference between short and long vowels in Latin is primarily derived from the fact that poets used regular alternations of long and short syllables in their poetry. The difference between a short and a long [i] was not itself written down, but we can deduce it from the fact that certain words with <i> letters only occur in certain positions in the line (those positions which require a long vowel), whereas others occur only in other positions (those which require a short vowel).

nonsense words

A next step on the scale of artificiality is that the researcher invents their own data, in order to test their theories. These can be *nonsense words*, i.e. word-like sequences of consonants and/or vowels which do not occur in the lexicon. We can then test the differences between such sequences.

possible word

One of the best known examples of this methodology is due to Morris Halle (1923-2018), one of the most influential phonologists ever. He observed that there is a three-way difference between *brick, blick* and *bnick*. The first of these obviously is a regular English word, but the other two are not. Yet English speakers feel a difference: *blick* could be an English word which you just happen to not know, whereas *bnick* just can not be English, it is too different from English words we know. Halle called *blick* a *possible word*, and *bnick* an *impossible word*. Phonology is under this view not about the actual words (whether or not a sound sequence gets assigned a meaning is seen as a random fact) but about the set of *possible words*. If somebody invents a new machine and everybody calls it a *blick*, that does not really change English. If somebody could convince everybody to say *bnick*, that would lead to a small change.

One can go one step further and not just invent words but whole patterns. For instance you could wonder whether the pattern in (2) above could be reversed in some language — whether it would be possible to have a dialect of English with plurals like the following:

(5) a. -[z]: *fishz, passz*
b. -[əz]: *cates, parkes, lampes*
c. -[s]: *lambs, pans, beds*

We can test this by trying to teach people this dialect. If we have a (synchronic) explanation for the 'real' pattern, it might predict that the pattern in (5) is unlearnable, or will not survive in a community which tries to adopt it. The latter is of course a little bit more difficult to test experimentally, although it can still be approximated (for instance by having a first generation of experimental subjects teaching it to a next generation).

Interest in such artificial language experiments has grown considerably over the past decade. This particular paradigm carries over quite easily into the next:

Experimental evidence

Another step we can take is go beyond the impressionistic recording of the data, is to go to a laboratory and get experimental results. The advantage of this is obviously that we can make very precise recordings, test our hypotheses under highly controlled conditions, which might also be 'unnatural' but important to see how language behaves under such circumstances.

There are many different things one can do in the laboratory. I will divide these into two types of experimentation: phonetic and psycholinguistic. The boundaries between these two are not always clear, but scholars tend to identify themselves as either 'psycholinguist' or 'phonetician' and the research communities are more or less separate.

Interest in laboratory methods for phonology has grown considerably over the past twenty years. There are special conferences and a journal on 'Laboratory Phonology', but also in other conferences and in other journals there is a rising interest in seeing phonological theories be supported or falsified by well-designed experiments.

Phonetic measurement

Over the course of the past 100 years, phoneticians have developed a large toolbox of instruments and techniques to study the way in which the human body produces and perceives speech sounds as well as the way in which such sounds are transmitted from one person to the next. Over the past few decades, several (free) software packages have been developed which can be installed on any laptop so that it is very easy to install a private phonetic laboratory. If you have a good microphone attached to it as well, you can do analyses at a very high level.

There are several advantages to phonetic data, also for the phonologist. One is that our instruments can measure fine-grained distinctions which are not always perceivable for the human ear. Or perhaps we should say: which we cannot raise to the level of consciousness. (If people really cannot hear a difference at all, one can wonder how the distinction can play a role in human communication.) If such patterns are systematic, they might be at least as important as those which we can obtain by other means.

Another potential advantage is that phonetic measurements are performed with computers and therefore less dependent on human interpretation. If you study somebody's speech by just listening to it, you might be tempted, even subconsciously, to hear things which are not really there, just because your theory makes you expect them to be there. A computer cannot be fooled in that way. Related to this is the fact that phonetic measurements are typically more easily *replicated* by other researchers than for instance native speaker judgements: if you carefully describe the way you have performed your experiments, another researcher will be able to do it in the same way, and arrive at the same results.

These advantages do not mean that all results which are obtained by other methods are therefore worthless or unscientific. Many patterns in human language are quite clear and obvious to any native speaker. The fact that the plural suffix is *-z* in *beds* but *-s* in *cats* can be easily observed without any technological means by any native speaker of English, nor will there be a lot of intersubjective

disagreement about this. For the latter reason, such data are *intersubjective*: the subjective intuitions about them agree to a large extent. To many phonologists it would seem a waste of time and other means to try to establish such truths also 'objectively', using computers.

This is then one disadvantage of phonetic measurements: they are time-consuming, you need to put informants into sometimes uncomfortable circumstances, and you need special equipment (even if this is only a recorder and a laptop), and it is not always clear that these costs are worth the result. Furthermore, it is not always even possible to acquire phonetic data, e.g. when we want to study the phonology of an exctinct language.

Another disadvantage is that phonetic measurements are necessarily 'superficial': they can only measure things which are present outside the human mind or human society: acoustic signals, movements of the body. Even if we also consider brain scanning techniques — which are currently too expensive for most linguistic researchers, but will probably be used more and more in the not too distant future — we can see pictures of which areas of the brain are active at some point, but not necessarily about what this means for the human mind. Like in the case of the corpora above, we cannot always see what the speaker intends, only what they actually do. If we want to know such things we have to apply techniques from psychology, as I will discuss in the next section.

Psycholinguistic experimenting

Psycholinguistics

Psycholinguistics is the field at the intersection of psychology and linguistics: it studies the way in which humans process, produce and acquire natural language. The results of this type of research are of course particularly relevant if one takes the psychological point of view and sees language as primarily something which belongs to the human mind. (To be precise, one sometimes distinguishes beteen *psycholinguistics* which is the linguistic discipline at the interface and *language psychology*, which is a branch of psychology.)

A typical psycholinguistic experiment has a number of speakers of a certain language perform a relatively easy language-related task. For instance, they listen to a number of words and press a button when they recognize that word as belonging to their language. Their performance is then measured in various ways, e.g.: how long does it take for them to press the button? How many mistakes do they make? And how are these factors influenced by others, e.g. the fact that they have just heard a word which sounds very similar in some way (starts with the same sound, has the same number of syllables)?

In this way, the psycholinguist hopes to find out how language is represented in the brain. In our example, which dimensions of sound similarity count as relevant for finding a word in the lexicon we as speakers all have in our head? In turn, such knowledge about the internal structure of the lexicon might teach the phonologist something about how sounds are organized in language and, inversely, phonologists' insights should ideally guide the psycholinguists' research.

A special place has always been occupied also by the study of (first) language acquisition. The language learning child to some extent has to face the same task as the linguist: she has to figure out what the system of her language is. There is one important difference, which is that the child apparently somehow knows how to go about this task and accomplishes it within a few years, while linguists continue to puzzle over the details. By following children closely, we can try to learn from the way in which they apparently acquire all these data.

But there are several other reasons why acquisition is important. One of these is that the fact that every generation has to learn the language from their parents is probably an important factor in language change. Small things can and do go 'wrong' in that children sometimes construct a slightly different language based on the input of their parents. If we are thus interested in the details of diachronic phonology, it is crucial to understand how such acquisition works. (Since language change may also be caused by adults having to learn a foreign language, studying second language acquisition can be relevant for the same reason.)

Formal evidence

A final type of evidence for a specific theory comes from properties of the theory itself. Scientific theories are generally supposed to be more successful if they are *elegant* and *restrictive*. Since these are properties of scientific theories in general, and not specific to phonology, we will not discuss them in detail here. There are many guides on scientific practice and the philosophy of science which propose are applicable and can help us find a good way of finding truth in our small subdomain of reality.

Restrictiveness

Restrictiveness of a theory refers to the number of things which are impossible according to the theory. The optimal theory is one which is applicable to every real object in the world, but *not* to anything else. The best theory of phonology would for instance describe and explain exactly all the sound sytems of languages in the world, but not arbitrary collections of sounds which can never be part of a real language. It should not just explain why the English plural system in (2) exists, but also why the fake sytem in (5) does not.

The arguments for formal types of evidence tend to be philosophical. An important argument why theories should be restrictive is that in this way they are more falsifiable. A theory which can account for everything, cannot account for anything at all. This is an objection which is for instance sometimes raised against current 'large language models' such as ChatGPT as models of human linguistic behaviour: such models seem to be able to deal with many kinds of patterns, including those that are not attested in human language and seem unlearnable for humans.

A second kind of formal criterion is *elegance*. If we have two theories of one and the same phenomenon, we prefer the one which is more elegant. The problem with this obviously is that it is an aesthetic criterion which thus may be subject to personal preferences and which seems difficult to define in an objective manner. However, there are several well-known and widely accepted principles of elegance which lead to at least some amount of intersubjectivity. The most famous among these no doubt is *Occam's razor*, named after the fourteenth Century English scholar William of Ockham. The principle says that a theory should not contain any unnecessary assumptions (such assumptions should be thrown out): the theory with the smallest set of assumptions is preferable.

Occam's razor

That in turn implies one should always go for the simplest theory, the one with the smallest number of assumptions. It is usually very difficult to really compare theories, because it is difficult to count assumptions, but everybody would agree that a theory which would require a desciption of 5,000 lines for a set of data is less simple than one which can be described in 5 lines and accounts for the same data. The latter therefore is preferable, all else being equal.

This is true in particular if the descriptions are written in the same language, which preferably should be highly formalized, like a mathematical notation, or a computer language. It has become more and more popular over the past decades to build computer models which mimic the way language works according to a certain theory (that is thus different from the large language models, which do not describe a particular theory). The goal of such computer implementations is not only to show that a theory is indeed really elegant, but also that it works in the first place. It may not always be easy to check all the predictions of a theory by hand. If you have written a computer program which does things exactly as you say it does and of which the outcomes are furthermore as they are in the real world, you know that the program contains the whole theory (there are no hidden assumptions) and furthermore that it works as required.

So which type of evidence is the best?

We have now given an overview of many types of data that are available to the phonologist. The overview may not be comprehensive, but most data that a phonologist deals with will be of one of the types described here.

There is an understandable human tendency to reduce such an embarassment of riches to something more tractable and declare that only some of these data types make sense, that the others are not acquired by the right methodology and that the rest can or even should be ignored. The linguistic literature, also that on phonology, is rife with such arguments that we should restrict ourselves to one particular set of data.

This is not the approach I take in this textbook. Even if you believe that certain types of data are inherently more useful than others, all of them are used in phonological argumentation, so a student of phonology should be able to deal with them and evaluate them. Furthermore, I believe that evidence for the kinds of patterns we study in phonology can be found in all of these data, although we should always be careful: each of these type of data has their problems and might be 'polluted' by other factors.

1.3 Phonological theory

Phonologists thus deal with a large variety of data, and try to find patterns in them. These patterns are then described in a theory. It is the ultimate goal of linguistics to describe what is a possible human language and what is not. Phonology shares this goal as far as the organization of sound patterns goes. There is a second goal, which is to fully describe the sound patterns which we find in the existing individual languages of the world; some phonologists work for instance exclusively on the phonology of French. In both cases, we are dealing with phenomena at a higher level of abstraction, so that we need a theory to describe them. Formation of such theories is an important goal of any scientific endeavor.

These two activities mutually feed each other. One cannot pretend to study what languages have in common, or what makes them different, without having a detailed knowledge of individual languages. On the other hand, in studying the phonology of such a language, it is important to know which aspects of the language are familiar also from other languages and which are unique for the language in question. The theory of French feeds the theory of language, and vice versa.

I have already pointed out that an important aspect of any serious scholarly discipline is that it is *cumulative*: we always try to build on the work of other scholars. This is not just true for building on other people's data, but also for the insights that have been embedded into theories. There are many questions, there are many mysteries. Other people will have (had) something to say which sheds light on these questions and mysteries. It is very inefficient if every individual starts all over again every time we study a subject matter. This is also how things are with phonology; this forms the justification for textbooks like the current one, where we try to synthesize in a didactic format what the current community of phonologists tends to see as the truth on the sound structure of human language.

cumulative

A brief historical sketch

Phonology arguably was one of the first domains of grammar that linguists got a grip on. In the 19th Century, one did not yet distinguish between phonology and phonetics: both terms were in use to refer to all kinds of phenomena related to the sounds of language. But no doubt the most prominent discovery of 19th Century linguistics was that one could reconstruct a thousands-year old language called Proto-Indo-European, based on careful comparison of languages from a large strip of territory from Europe to India — the area to which this language had apparently be spread. This comparison was initially based almost exclusively on sounds. It was shown that by comparing the phonological structure of words in Latin, Greek, Sanskrit and many other languages, one could get to a reasonable idea of what the words must have looked like in a language from which these other languages originated. This work was seen by many as the first example of successfully applying scientific methodology to an empirical topic that had always been seen as belonging to the humanities.

The discipline of synchronic phonology started as a recognizable scientific enterprise somewhere at the beginning of the twentieth century. A reasonable starting point is the *Course on General Linguistics* by the Swiss linguist Ferdinand de Saussure (1857-1913). In this book, Saussure introduced many concepts which turned out to be foundational for modern linguistics. The distinction between synchronic and diachronic language study, for instance, is due to him. But one of his key ideas was that languages can be studied as coherent systems: that for instance the sounds of a language exist in patterns. For the study of sounds, this implied the introduction of a distinction between phonetics and phonology, where the latter was uniquely occupied with the (linguistically relevant) patterns.

In terms of the discussion in this Chapter, Saussure had a *sociological* view on these patterns: he believed that 'language' only exists as a property of a community. The speech of each individual only gives an imperfect reflection of the more perfect abstract object which everybody shares. This was also the view of *structuralism*, the most influential linguistic paradigm both in the United States and in Europe until the 1950s, of which Saussure is usually considered to be the father. Although there was also some interest in syntax and morphology, the most succesful branch of structuralist linguistics was phonology. Important scholars like Leonard Bloomfield (1887-1949), Nikolay Trubetzkoy (1890-1936) and Roman Jakobson (1896-1982) produced very influential work laying the theoretical foundations of the field. In particular, Bloomfield's book *Language* and Trubetzkoy's *Foundations of Phonology* are still important references for any (advanced) student who wants to understand what phonology is about.

structuralism

The advent of *generative linguistics* (of which the main phonological event was the appearance of *The Sound Pattern of English* in 1968 by Morris Halle and Noam Chomsky) brought about several changes. The most important one was a change in orientation: generative phonology puts the reality of phonological patterns unequivocally in the human mind. At the same time, Halle and Chomsky brought several technical innovations and improvements to the theory, although there is also a lot of continuity (and cumulativity) in technical aspects of the theory.

generative linguistics

In the 50+ years following the publication of the *Sound Pattern*, the theory has developed in many ways, basically to the point of having become unrecognizable. If you have studied this book, for instance, you will find it much easier to follow contemporary literature than the work of the 1970s.

A striking aspect of the current phonological community is that it is very pluriform: in a given phonological conference you will find many different approaches. Although most phonologists probably still take a psychological view of phonology, the sociological view has also again gained ground. Yet although there are a lot of debates about many aspects, as is the case in any healthy discipline, there is also a large common ground — a vocabulary in which scholars can communicate about their ideas. My aim in this book is to bring you up to date with that vocabulary, showing some of the diversity, but focusing on what I consider to be the hardcore — the theoretical body details of which are sometimes questioned and often debated, but that in the end still is known by people who call themselves phonologists.

1.4 Structure of this book

In this book, we will start with the smallest constituting elements of language. As we will see in Chapter 2, these are not consonants and vowels, but even smaller units which can be combined in certain ways to form units that are similar to such sound units, and Chapter 3 looks into the methododology for finding out how phonological structures are organized in a given language.

From there, we will start looking at larger and larger units organizing sound in human language. Chapter 4 discusses autosegmental theory, which explains how these 'atoms of language' can be combined into larger wholes, Chapter 5 then explains the evidence for assuming that consonants and vowels (together called *segments*) do not only have an internal structure, but are in turn also organized into larger units, viz. *syllables*, in presumably all languages of the world.

The topic of syllables also introduces the notion of *resyllabification*: syllable structure (and other phonological structure) sometimes changes, for instance when two morphemes are put together in a word. We have very precise ideas about the properties of the computation that is necessary to make such changes happen. These are discussed in Chapter 6.

Chapter 7 builds on the introduction of syllables in Chapter 5 and discusses how syllables are grouped into metrical feet, which are used among other things for word stress. Chapter 8 then shows that there are even higher levels of phonological organization, leading all the way up to constituents that roughly correspond to the sentence and even the whole utterance.

The last two chapters discuss the way in which phonology relates to other domains. Chapter **??** deals with the relation with morphosyntax, and Chapter **??** discusses several applications: in orthography, in speech recognition and in our understanding of Sign Language.

Every Chapter contains a section with 20 exercises. (Only this first chapter, which is very general, has only 10.) It is very important to try to solve those exercises; there is only one way of learning to do phonology: by practicing. Some of the exercises can be answered by just using the material which is presented in the book, but for others you will need to collect data, online or in other ways. I will always explain where you can find such data. More information can be found in the website which accompanies this book [UNAVAILABLE]. You can also find the answers to the exercises there, as well as a more extended (and updated) version of the reference sections which are at the end of every Chapter, and which you can use if you want to know more about a certain topic.

1.5 Exercises

1. Would the study of the following topics belong to the field of phonology? Discuss. If you think a certain topic is not studied by phonologists, which other field does it belong to?

   (Ła) The Hawaiian language has eight consonants: /p, k, ʔ, h, m, n, l, v/.

   (Łb) In French, some adjectives end in a consonant in the feminine, but not in the masculine. ('Small' is [ptit] in the feminine, and [pti] in the masculin; 'good' is [bon] feminine and [bõ] masculine).

   (Łc) In some cultures, homosexual men speak differently than heterosexual men. It can be shown in experiments that people are sensitive to these differences and can tell above chance level what the sexual orientation of the speaker is.

   (Łd) The English [b] sound is much more acoustically similar to the French [p] than to the French [b].

   (Łe) There is no language in which every prime numbered syllable starts with a *p*; furthermore, people cannot learn such a 'language'.

   (Łf) Leaving out a final *d* or *t*, as happens in some dialects of English (*I kep*) is considered incorrect.

(Łg) In some dialects of English, the /r/-sound is not pronounced after some vowels (like in *car, mother, more*) except if the following word starts with a consonant.

2. For each of the topics mentioned in the previous exercise, explain which kinds of methodology could be applied to shed light on the issues involved.
3. For each of the following strings, decide whether they are existing words of English, possible words, or impossible words: *blobber, brankal, rooytkd, trapeal, blistras, topiq*.
4. Explain the difference between synchrony and diachrony in language. Why does holding a Platonic or a cognitive view of language usually imply more interest in synchronic data?
5. There seems to be an intimate relation between having a Platonic view of language and being interested in formal types of evidence. Explain.
6. *Duality of patterning* is supposed to be a defining property of human language. Explain why the following systems lack this property:

   (Ła) Programming languages like Java, Python, C++.

   (Łb) Animal communication systems, such as primate calls.

   (Łc) Western classical music.

7. Suppose we want to study the phonology of Sanskrit, an Indo-European language that has been extinct for a long time, but of which we have a lot of written record, like poems and prose. What kinds of methods could we use to study this topic?
8. For each of the major types of evidence in section 1.2, give an example of how they could be applied also to morphology or syntax.
9. 'All phonological research should be based on data that are acquired through phonetic experimentation.' Comment.
10. What are the advantages of building a computer model of a theory? Can you also think of disadvantages?

Sources and further reading

Section 1.1. There are currently several other textbooks on phonology in the market; implicitly, they all take a cognitive view on the location of phonology. The most well-known are Gussenhoven and Jacobs (2005); Odden (2005); Hayes (2009). They each have their merits, and in case you somehow struggle with a topic, it can always help to see how somebody else explains it. For more extensive information you can also refer to recent handbooks such as van Oostendorp et al. (2011); Goldsmith et al. (2014); de Lacy (2007).

It is also convenient to have some basic knowledge about phonetics. There might be even more of these than there are introductions to phonology. One I like is

A well-known general textbook about linguistics is Fromkin et al. (2018); an open access alternative is Anderson (2018). Some good syntactic textbooks are Radford (2012); Adger (2003); van Gelderen (2017), and two good books on morphology are **?**Booij (2007). Nice linguistic introductions to phonetics are Ashby and Maidment (2005); Knight (2012)

The Kalkatungu data in this section are from Blake (1969).

Section 1.2. There are unfortunately no general introductions into phonological methodology, describing all the different types of data and data collection, but in the chapters that follow we will see individual cases. A recent book on fieldwork is Sakel and Everett (2012).

The International Phonetic Alphabet (IPA) is maintained by the International Phonetic Association which (among other things) has a chart of all the IPA symbols on its website. The IPA symbols are part of the standard Unicode character set, which implies they can be used on any reasonably modern computer system; how to do that is explained on the website of SIL International.

A phonetic app that is often used is Praat, developed at the University of Amsterdam in The Netherlands by Paul Boersma and David Weenink. Praat is free of charge and available for all major platforms; it is also regularly updated. Next to a popular module for making acoustic measurements, it also contains modules for applying statistic calculations, modelling language acquisition and many other tasks which a computer might want to perform for a phonetician and/or a phonologist.

Section 1.3. Anderson (1985a) gives an authoritative overview of the history of phonology until approximately 1980. Full references to the books mentioned in the main text are Trubetzkoy (1939) and Chomsky and Halle (1968).

The atoms of language

2.1 Segments

An important activity of linguists is the pursuit of *linguistic universals* — properties which all languages have in common. The reason for undertaking this enterprise is quite obvious: if we know what all languages share, we can say that we have clearly discovered something essential about human language. Furthermore, every claim about linguistic universals can be falsified, and this is a desirable property for any scientific theory, as it means that what we are saying is not trivially true. **linguistic universals**

Once we have found a phonological universal, there are obviously many possible explanations for it. It might be that something is universal because of universal properties of the human body (for instance the universal that no language uses the sound one can make by sticking one's tongue in one's eye), or of human cognition (for instance the fact that no language has words with e.g. 5,000 different vowels and consonants which would be hard to remember), or to the fact that languages are used for communication (e.g. that vowels and consonants in a language tend to not resemble each other so closely that mistakes are very likely). After we have found the universals, we still have to find the explanation. But finding the typological generalisations comes first.

One universal is that it is possible in all languages to divide the sound stream into sequences of vowels and consonants, together called *segments*, and that the set of members this set ranges somewhere between 15 and 150. These are actually two different, but related universals: **segments**

(6) a. The sound inventory of all languages can be divided into a finite set of segments, $\mathscr{S}$.
    b. In all languages, $\mathscr{S}$ can be subdivided into (complementary) subsets of vowels and consonants.

'*Segments*' is the technical term for individual speech sounds. If you have been trained from an early age in a culture which uses an alphabetic system to write its language, this may sound trivial, since alphabetic writing is based on the possibility of (6). People tend to think, even if subconsciously, of written language as primary, and to consider it obvious that every language can be written using a finite set of letter symbols, and furthermore that some are called vowels and others consonants. **Segments**

We should however be cautious that we are not misled by this cultural invention to think that there is something real about these distinctions also in spoken language. For this reason, it is always good to be critical and see what independent evidence we have for the reality of $\mathscr{S}$, and in particular for the claim that it can always be subdivided into a subset of vowels and a subset of consonants.

Evidence from writing systems

The fact that an alphabetic writing system can work for all languages, and that also languages which use a different writing system can be transcribed at least in the extended symbol set of the International Phonetic Alphabet, is not a triviality, and is itself a piece of evidence that something is right about (6).

As I already mentioned in the beginning of this book, would be possible to construct a language in which every morpheme corresponds to its own unique sound: *man* is designated by a sneeze, *woman* by clapping your hands once, and *love* by flapping your lips. In such a language *a man loves a woman* would thus sound as SNEEZE-FLAP-CLAP. Such a language would not necessarily be difficult to produce or to perceive by humans. The fact that no human language has such a lexicon, but that instead words are always built out of a relatively small inventory of segments, is thus meaningful.

This implies that alphabetic writing is a technology that builds on a cognitive structure which seems inherent in human language. We have to keep in mind, though, that the relation between sound and letter can be a very complex one. One and the same sound can be represented with many different letters and letter combinations in a language like English, as *truly, do, shoe, soon, true, lawsuit, routine, two, screwed, jewel, manoeuvre, rendezvous, throughout* and *coups* show: all have an [u] sound. The reason for this often is that spelling does not just represent the current sound structure, but also the history of the language. In other languages (for instance, in Italian) the relation is more one on one.

The difference between consonants and vowels becomes apparent in spelling. There are many examples showing that a sentence without vowels can still be read, whereas a sentence without consonants or leaving out a random set of segments, usually cannot:

(7) a. *Mst ppl wll ndrstnd ths sntnc.* (Most people will understand this sentence.)
b. *U iuay ooy a uea i.* (But virtually nobody can understand this.)
c. *Ms epe il o udrtd hs ihr* (Most people will nog understand this either.)

Related to this is the finding that when speakers hear the nonsense word [kebra], they are more likely to think of a *cobra* than of a *zebra*: the former is one vowel 'away' from the word they have just heard, whereas the other is one consonant away. Changing a vowel thus seems to matter less than changing a consonant. (This experiment was done with speakers of Spanish and Dutch, which are both Indo-European; we thus have to take some caution, also because both languages are written in a — Latin — alphabet.)

Computer-mediated communication, for instance by WhatsApp and Instagram also sometimes uses this technique of leaving out vowels to shorten a message (*txtng*). Semitic languages such as Hebrew and Arabic use this property even systematically in their 'official' writing systems: vowels are not usually written in their script, except sometimes in special versions for people learning to read and write. There are no languages in the world which do the opposite, i.e. which write only the vowels of words but not the consonants.

One reason for this asymmetry may be that there are as far as we know always more consonants than vowels in the segment inventories of languages (for instance: Italian has 24 consonants and 7 vowels, Arabic has 29 consonants and 3 vowels, Malay has 20 consonants and 5 vowels). This example thus not only shows that there is a distinction between consonants and vowels, but also that this distinction involves an asymmetry between the two sets. Again there must be some explanation for this assymetry, since there are languages which have tens of contrastive vowels: why aren't those accompanied with relatively small sets of consonants?

Semitic templates

Semitic languages provide another well-known type of evidence for the existence of consonants and vowels beyond their writing system. You can see this in the following examples from Modern Hebrew (Semitic, Israel):

(8)	gadal 'to grow'	higdil 'to enlarge'	gdila 'growing'
	ʃatak 'to keep quiet'	hiʃtik 'to quieten'	ʃtika 'silence'
	sagar 'to close'	hisgir 'to extradite'	sgira 'closing'

If you study these examples closely, you will observe that the three words in every line have several things in common. This is true both for the meaning and for the form. As to the meaning, the first three words are about *size*, the second set of three about *silence*, and the third set about *being closed*. As to the shape, the first three words all contain the consonants *g, d, l* in that order, the second triple have *ʃ, t, k*, and the third triple *s, g, r*.

In Semitic philology, it is usually assumed that these general meanings are indeed attached to the consonants (and they are called *consonantal roots*). **consonantal roots** If you look further, you will discover that the vowels also have something to add to the meaning, for the *columns* in the table also share certain things: the first column contains simple verbs, the second column contains so-called *causatives* describing how something is made larger, quieter of closer, and the **causatives** third column has nominals. It is no accident that the vowels are also the same in each column. We thus get the appropriate meaning of a word by combining the meaning of the consonantal roots with those of the vocalic patterns.

Templates of these type are found in all Semitic languages, but also in unrelated languages such as Yowlumne (Yakuts, North America) and possibly Rotuman (Oceanic). This is thus some kind of mechanism that people can apparently use to form their morphology.

Psycholinguistic evidence

There is also quite some evidence that consonants and vowels are represented in different parts of the brain. For instance, it has been shown that patients suffering from *aphasia* (which is a language disorder, usually caused by head **aphasia** injury or a stroke) can sometimes be affected only in the vowels — making more errors producing them or keeping them apart — while other patients have more problems with the consonants.

It has also been suggested that listeners tend to pay much closer attention to small oscillations in the production of vowels than that of consonants. The reason for this might be that variation in vowels is more informative, for instance telling us whether the word is emphasized, what kind of intonation it carries, or even what the emotional state of the speaker is.

Finally, also brain scanning has shown differences between consonants and vowels. For example, changing vowels relative to consonants while a subject in a brain scanner was reading aloud words, increased activation in a right middle temporal area, whereas changing consonants relative to vowels increased activation in a right middle frontal area. There thus is evidence that consonants and vowels might even be physiologically distinct in the human brain.

2.2 Contrast and the feature

features

Even though we have quite some evidence for their existence, segments are not the ultimate primitives of phonological theory. It is usually assumed that there are smaller building blocks — called *features* — which together somehow form the segment.

We can see these features as instructions for the articulatory organs. Take the French consonant [b]. In order to pronounce this sound, we have to close our lips and to vibrate our vocal folds (among other things). These are two instructions, one to the lips, and one to the vocal folds. These instructions correspond to features, [Labial] (*labia* is Latin for lips) and [Voice]. (Feature names are typically written between square brackets; the curly brackets indicate that we are dealing with a set of them.)

[b] shares these features with [v], but is different from the latter sound because it is pronounced with an explosion whereas [v] is pronounced with friction. This explosion also corresponds to a feature, [Stop]. The consonant [b] therefore is supposed to have roughly the structure of a set of features:

(9) [b] = { [Stop], [Labial], [Voice], . . . }

phonological representation

(9) is a *phonological representation*, a structure to give a form to our ideas of what a segment such as [b] is. We are assuming that such a [b] is thus a set of things, and those things are features. A representation like this might be translatable in various way to the physical world. For instance, the features might correspond to certain characteristics of the speech signal or, as I suggested above, be seen as instructions to the speech organs. We will refine our representations in the rest of this book, but there ultimate building blocks will remain to be features.

I put the dots in this representation, because possibly there are other features which a [b] has. However, not every possible movement we make with our lips, or anything else we can observe about the sound counts as a feature. For instance, when pronouncing the [b], a speaker has to put her tongue in a certain position — typically somewhere low in the mouth — just because any physical object has to be somewhere in space. However, [Low] (for a low tongue position) is not assumed to be a feature of [b], even though it *is* a feature of a vowel such as [a]. The reason for this is that there are other vowels which do not have such a low tongue position, such as [e] or [u], but there is no difference between a 'low' [p] and a 'non-low' [p]: the difference would be inaudible, so it would not serve any function in communication.

How do we know which properties get a formal status in the theory as features? We usually assume that only those traits that are *linguistically relevant* count as features. The primary type of linguistic relevance for features is *distinctiveness*. The feature [Voice] is distinctive on French [b], because there is another sound in the language which has exactly the same features, except that it is not voiced: the [p] is also a labial stop, but the vocal folds do not vibrate. On the other hand, [b] is not distinguished from any other sound only by the position of the tongue; French does not have a consonant with closed lips and a tongue in a high position (nor does any other language as far as I know, because the position of the tongue is hard to hear if the mouth is closed).

linguistically relevant

distinctiveness

The traditional test in phonology — and possibly the oldest type of phonological methodology — to decide that two sounds are different is the so called *minimal pair test*: we construct two words with different meanings which are only different in the sound in question. A relevant minimal pair in French would be:

minimal pair test

(10) *pont* 'bridge' [pɔ̃] : *bon* 'good' [bɔ̃]

Since the two words have different meanings in French, the difference between [b] and [p] — the only difference we can find between these words — must have a relevance in French. Since these sounds are only distinguished by voicing, this means that French has a feature [Voice]. The ultimate building blocks of French will thus be a small set of these features. (Notice by the way that there are more differences in the orthography, as the word for 'bridge' ends in a *t*; this is however considered irrelevant for phonology, since that *t* is not pronounced.)

There are languages which have no distinctive voicing at all. Such languages do not distinguish between [p] and [b] or [k] and [g] at all. This typically means that speakers will choose the form that is easiest from the point of view of articulation (for labials this is the voiced form and for velars the voiceless form). But speakers are also free to vary, on an individual basis, or because in certain phonetic contexts the other version might still be easier, or sociophonetically. This means that in such languages e.g. both [p] and [b] can occur, but they have the same phonological representation.

In the remainder of this section, I will briefly discuss some important types of evidence that have been adduced for the existence of the phonological feature.

Phonological activity

One important type of evidence comes from phonological alternations (the topic of Chapter 3). Consider the following examples from Turkish (Turkic, Turkey), paying particular attention to the final consonant of the stem:

(11)	kalıp	'mold'-NOM	kalıb-a	'mold'-DAT
	kap	'container'-NOM	kab-a	'container'-DAT
	kanat	'wing'-NOM	kanad-a	'wing'-DAT
	tat	'taste'-NOM	tad-a	'taste'-DAT
	güve[tʃ]	'clay pot'-NOM	güve[dʒ]-e	'clay pot'-DAT

You can see that the final consonant is different when it appears at the end of the word (such as in the Nominative) than when it appears before a vowel (as in the Dative). In the former context we find { p, t, tS, }; in the latter { b, d, dZ }. The difference between these sets is that the latter all have a feature [Voice], which the former lack:

(12)	[p]	{ [Labial], [Stop] }	[b]	{ [Labial], [Stop], [Voice] }
	{ [t]	[Coronal], [Stop] }	[d]	{ [Coronal], [Stop], [Voice] }
	[tʃ]	{ [Palatal], [Stop] }	[dʒ]	{ [Palatal], [Stop], [Voice] }

(I introduce the features [Coronal] and [Palatal] here. The former denotes sounds made with the tip of the tongue at the front of the mouth, the latter sounds which are made slightly more to the back. We will return to these features in the next section .) We can thus draw a generalization over the data in (11) as follows:

See Section 2.3

(13) When the Dative has features $F_1, F_2, \ldots, F_n$, as the last segment of the stem, the Nominative has features $F_1, F_2, \ldots, F_n$, except for [Voice].

This is more elegant than simply stating for every word what the two forms are, but it also more elegant than stating that words which have [b] in the Dative, get [p] in the Nominative, while those with a [d] get a [t]. We generalize over all these cases, and the elegance which ensues from this is what we aim for in constructing a theory.

sonorants

Interestingly, consonants such as {m, n, r,...} (called *sonorants*) do not participate in this kind of alternation:

(14)	adam	'man'-NOM	adama-a	'man'-DAT
	tavan	'ceiling'-NOM	tavan-a	'ceiling'-DAT
	zar	'die'-NOM	zar-a	'die'-DAT

From a phonetic point of view these sonorant sounds are voiced in Turkish, just as they are in English and many other languages of the world. There are no counterparts which are not voiced, and in particular there their voicing they are not voiced. Given the logic just explains, this means that in the phonology these sounds would not have a feature [Voice]. This explains why they do not participate in this asymmetry: the set of features is the same in the Nominative as in the Dative.

Symmetry of inventories

Another argument in favour of seeing features as the building blocks of phonological representations, is that we predict that inventories of segments tend to be symmetric, and this is indeed what we find.

Suppose that the consonant inventory of a language consists of the features [Labial], [Stop] and [Nasal]. You can check that these features can be combined in 6 different ways ($2x2x2-2$), if we also allow a consonant without any features, but assume that the combination of [Stop] and [Nasal], which give conflicting instructions, is not allowed. This describes almost exactly the consonant inventory of Hawaiian:

(15) Hawaiian (Poylnesian, Hawai'i)

[Labial]	∅	
p	k	[Stop]
m	n	[Nasal]
w	l	∅

We thus have a very regular pattern: two stops, two nasals, and two non-stop, non-nasal consonants. And within each of these pairs, one is labial whereas the other is not. This regularity is very nicely described with features. What is more, it is extremely common for languages to have their consonant inventory structured in such a regular way: typically they have the structure of rectangles. There are sometimes holes in these rectangles (combinations of features that are not realized in the language in question), but statistically if a language has the features F_1 and F_2, it is very likely to also have the sound combining F_1 and F_2.

Hawaiian has two further consonants ([ʔ] and [h], two sounds which are not articulated in the mouth) which do not fit in this pattern. That is also a very common property of languages: there is a tendency towards symmetry, but this is rarely absolute. There will be some gaps in the table or some sounds which do not really fit.

Any theory will have to be able to take into account both of these factors: that there is a universal tendency towards having a symmetric inventory of sounds, and that very few languages have an inventory that is absolutely symmetric. One way to do it is to assume that the 'sound grammar' of languages consist of a set of features plus rules about how these features can be combined — one rule which is necessary for Hawaiian is that [Stop] and [Nasal] cannot occur together in a segment, as we have seen.

Language acquisition

Features may also play a role in language acquisition: there is evidence that children acquire features rather than segments. Young children (roughly in their first year of life) acquire the individual sounds of their language. This process is very complicated, and I cannot go into it in full detail here. But roughly, after a first stage in which they produce all kinds of sounds and learn to say a few words like *mummy*, which do not yet seem to be analysed in segments, let alone in features, they start building up the system. And they seem to use features to do so.

Typically, a French-speaking child, for instance, may at some point have acquired the consonants *p*, *t* and *k*. These are all three voiceless stops, and they only differ from each other because the first is [Labial], the second [Coronal] and the third [Dorsal] (pronounced with the back of the tongue). These will then be the only phonological features that the child has acquired.

A next step might be that they acquire the feature [Voice]. Our prediction is then that they will be able to combine this new feature with the ones they already have. In other words, that they will acquire *b, d* and *g* at more or less the same time, combining each of the three existing features with the new [Voice]. Similarly, at some point she will learn *m, n* and ŋ at the same time, since these require combining a new feature [Nasal] with the three features that are already in place.

Just like with the preceding evidence, feature theory alone does not make a perfect prediction. Children each follow their individual path and are influenced by all kinds of other factors, such as how often they hear a sound (a child called *Robin* might be quicker in learning the [r] sound). But as far as we know, features do play an important organizing role in the acquisitional path.

Speech errors

Speech errors also provide a possible source of evidence. While we are speaking, we constantly make small errors. We usually do not really notice them in our selves or in others, even when we quickly correct them — and the latter does not seem to happen very often either. But when you pay special attention you can notice them and by studying them carefully, we may learn many things about phonological structure.

Phonologists have observed, for instance, that errors often involve exchanging two consonants or vowels in a word or word group:

(16) a. *fish grotto → frish gotto*
b. *fresh clear water → flesh queer water*
c. *brake fluid → blake fruid*
d. *add hoc → odd hack*
e. *fish and tackle → fash and tickle*

In the examples in (16a)-(16c), two consonants have traded places; the same has happened to vowels in (16d)-(16e). Interestingly, it appears very difficult, if at all possible, to find examples where a vowel has exchanged positions with a consonant. This is yet another indication that vowels and consonants somehow exist in different dimensions, at least in their representation in the human brain.

These data also give evidence that segments can indeed sometimes be isolated: in (16a) it isn't the whole consonantal group *gr* that is pronounced in the wrong position — although such errors do also exist — but just the [r]. In order to make such a mistake, the speaker must thus be able to somehow separate this sound from the others.

Also individual features can sometimes move around within a word. I have a colleague whose child insisted that she wanted to eat *skabetti*. If you analyse this and compare it to the intended *spaghetti* ([spageti], disregarding the length of the consonant [t]), you will observe that the place in the mouth in which the first two plosives are pronounced have been interchanged. However, the first obstruent is still voiceless ([k], not [g]) and the second one voiced ([b], not [p]).

This type of child language behaviour may strictly speaking not count as a speech error, but we find similar mistakes also in 'real' speech errors of adults:

(17) a. *Cedars of Lebanon* → *Cedars of Lemadon*
b. *pity the new teacher* → *mity the due teacher*

2.3 A universal feature set

We have thus seen that there is quite some evidence that segments are not the smallest possible units of linguistic analysis, but that they can be further decomposed into features. We have also mentioned some of these features: [Stop], [Nasal], [Labial] and a few others.

One could ask many questions about the status of these phonological features. Do all languages employ the same set — or at least a selection out of this universal set? Or do they each make up their own features? And what properties of sounds do they refer to? For instance, the feature label [Labial], which we have used so far, refers to a certain *articulatory gesture* which speakers make when producing the sound: they move their lips. **articulatory gesture**

There are logically speaking at least two alternative possibilities. The first is that features do not refer to articulatory properties, but rather to acoustic or perceptual ones, e.g. the effect which the sound has on the speech signal. This view is entertained also by many phonologists. We will briefly discuss one phonological theory in which this point of view is prominent in section 2.4 below.

The second alternative is that phonological features have no relation to the phonetic shape of sounds at all. This means that they are purely abstract; they would have the shape [α], [β], etc. Although this position has been suggested by phonologists (often using the term 'substance-free phonology'), even those people who accept prefer abstract labels, in their actual analysis usually use phonetically motivated ones for comprehensibility. The reason for this is that they usually do the job rather well, and furthermore are easier to keep apart in an analysis which uses many different features than abstract labels such as [F] or [α].

I will thus disregard these two alternatives and base our discussion on a version of articulatory-based feature theory, which seems to be still the *lingua franca* of feature theory. Even within that theory, there is some variation in terminology and labeling of the features. I have chosen the labels which I believe are most common, and have tried to be coherent, but it is unavoidable that you will sometimes encounter labels in the literature which are slightly different from the ones presented here. Using a dictionary or a phonetic lexicon will then usually help out.

The feature set is divided into a number of subsets which each describe the properties of sounds. We will discuss these sets in turn on the following pages.

Manner features

The first set of features describes the general way in which sounds are formed. These are called *manner features* or also *major class features*. We have seen one **manner features** **major class features** of those above, **[Stop]**, which denotes sounds that are produced with a temporary obstruction at some point in the vocal tract, and a subsequent burst when the outgoing airstream is released. Examples of stops (also called *plosives*) in English are [t, k, p, d, g] and [b]. I have already introduced the feature label **[Stop]** above.

plosives

Fricatives are sounds which are pronounced with slightly less obstruction. The result of this is that the outgoing airstream whirls and makes a fricated noise. Fricatives and plosives together are called *obstruents*, i.e. sounds in which the airstream is more or less obstructed during articulation. Fricatives are different from plosives in that the former carry the feature **[Continuant]**, and miss the feature [Stop].

Fricatives

obstruents

All consonants which are not obstruents are called *sonorants* (and the corresponding feature is **[Sonorant]**). In sonorants, the air can stream out of the lungs in a way that is more or less unimpeded, although it is still transformed in certain ways.

sonorants

The first class of sonorants consists of so-called *nasals* such as *m* and *n*. In these, there still is obstruction in the oral tract, but the airstream can escape through the nose. (You can check this by putting a little mirror under your nose while speaking; if you say a *p*, *b* or *f*, nothing happens. But when you pronounce an *m*, there will be some condensation on the mirror. This is the air that came out of your nose.) Nasals unsurprisingly have the feature **[Nasal]**.

nasals

The next class of sonorant consonants are the *liquids*, typically *r* and *l* (and in some languages also some other variants). In liquids, the stream goes really out of the mouth, but is deformed by the position of the tongue or, sometimes, the lips or the uvula. *R*-like sounds are distinguished from *l*-like sounds in that the former carry the feature **[Rhotic]**, indicating that there is some trill somewhere in the mouth (languages may vary in where this is realized) and the latter **[Lateral]**.

liquids

The final class of sonorants are the *glides*, such as English [j, w]. These are the consonants that are closest to the vowels. As a matter of fact, [j] is pronounced as a [i], and [w] as an [u]. The main difference is that the former are substantially shorter than the latter. It is often assumed that glides and vowels therefore have the same feature specification: they share the feature **[Vocalic]**, in contrast to all the other sounds we mentioned, which are assumed to carry **[Consonantal]**.

glides

There are several other manners of articulation of sounds, which English does not employ. For instance, some languages use *ingressive sounds*, which are produced while the air streams into the lungs rather than outside. Another different type of producing sounds is instantiated by *clicks*, which we find in languages of Southern Africa. These sounds are made not by air streaming in or out of the lungs but by sucking the tongue or lips and releasing them with a little explosion.

ingressive sounds

clicks

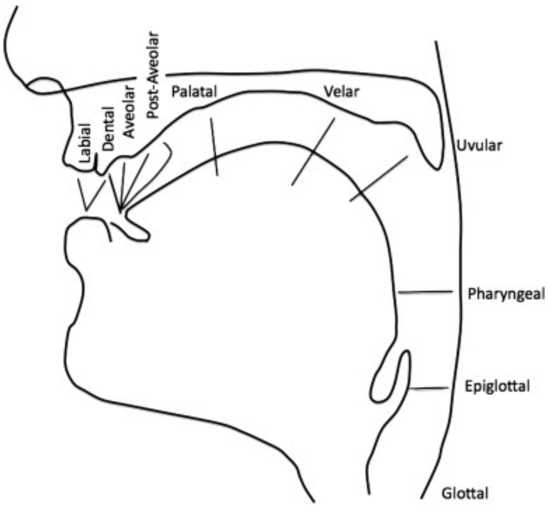
As we will see in Chapter 5, manner features are very useful in describing how vowels and consonants can be arranged in a word. For instance, an English word cannot start with a liquid followed by a plosive (**rtee*, **lpate*), although these clusters are allowed in reversed order (*tree*, *plate*). Many languages — although not all — have restrictions of these type, which often refer to the manner of articulation in some way.

Consonantal place features

The next set of features denote the place in the vocal tract where there is a constriction of the aistream in the case of obstruents, or where the airstream is deformed, in the case of sonorants: the *place of articulation*. **place of articulation**

The left-right dimension in the mouth does not play a role in the phonology of any language: it does not matter whether you put your tongue in your left cheek or your right cheek for instance, most likely because that difference is not audible. All relevant differences are placed in a line from the front of the mouth to the back of the throat, usually represented in a picture like this:

(18)



Ten different places of articulation have been distinguished in this picture. Within a very fine-grained phonetic analysis one might be able to distinguish between more, but these are definitely the most important ones for human languages. We will briefly discuss each of them in turn.

The first of these are the *labial* sounds, which are pronounced with (or at) the lips (*labia* is the Latin word for lips). Examples of these in English are *p, b, m* and *f*. The first three of these are *bilabial*: they are made by closing both lips (*bi-* means 'two'); the fricative however is *labiodental*: it is pronounced with the upper teeth on the lower lip. Labial sounds have the feature **[Labial]**. **labial** **bilabial** **labiodental**

The second broad class of sounds are called *coronal*: the corona is the small ridge just behind the teeth (the alveolar ridge), and these sounds are pronounced around that area. Sounds which are made on the ridge are called *alveaolar*; examples are [z, s, n] in English, but the language also has sounds which are pronounced with the tip of the tongue at the teeth, such as [ð, θ] (the initial sounds of *these* and *thing* respectively. Coronal sounds have the feature **[Coronal]**; where it is necessary to distinguish between different types, other features might play a role, such as **[Alveolar]** or **[Dental]**. **coronal** **alveaolar**

Putting the tongue a little bit further down the palate (but still on the hard palate), one produces *palatal* sounds. English [ʃ] (as in *she*) can be seen as an example. Palatal sounds are often still classified under the coronal sounds, but they might carry a (further) feature such as **[Palatal]**. **palatal**

velar
dorsal

One step further down the vocal tract, we find a class of sounds which are pronounced with the back of the tongue (the *dorsum*) at the soft palate (the *velum*). These sounds (such as *k*, *g*) are therefore called *velar* or alternatively *dorsal*, with a feature **[Dorsal]**.

uvular
pharyngeal
Epiglottic

At the far back of the mouth we find the uvula; some languages have sounds which are made here, and which are logically called *uvular* sounds. The French pronunciation of *r* ([ʀ]) is one of them. Within the throat we then find various sounds which are more exotic to the native speaker of English. Many language in Northern Africa have *pharyngeal* consonants, like the [ʕ] in Somali (Cushitic, Somalia). *Epiglottic* sounds like the plosive [ʡ] are found in Dahalo (Cushitic, Kenya). (There is discussion whether in the phonology we ever need to really distinguish between these pharyngeal and epiglottic sounds, but phonetically they are definitely different.) We can use feature labels like **[Uvular]**, **[Pharyngeal]** and **[Epiglottic]** corresponding to these.

glottal stop

Lowest down in the throat (at least as the production of sounds is concerned) is the glottis. We obviously use the vocal folds for every sound that is voiced, but the only consonant we call glottal is the *glottal stop* [ʔ], which you can hear, for instance, in Cockney English (the traditional dialect of London City) where it replaces other stops ([stɔʔ] instead of [stɔp]). Also *h* is sometimes seen as a glottal sound. The corresponding feature is **[Glottal]**, although it is also sometimes argued that glottal sounds simply have no place of articulation feature at all.

Laryngeal features

Every consonant by necessity needs to have a manner and a place. You can therefore safely assume that every consonant has at least two features. Next to this, a consonant can optionally have some other features.

laryngeal features

An important class among these are the *laryngeal features*. We already mentioned above that there is a difference between for instance a voiced *b* and a voiceless *p*, which can be expressed by assuming that the former has a feature **[Voice]** that is missing in the latter.

Many languages have these kinds of pairs (*b-p, d-t, g-k, v-f, z-s, D-T, G-x*, etc.) Some languages also have triples. Korean (isolate, Korea) is probably the most well-known example of this. Here are some minimal triples:

(19)	*bang* 'bread'	*pang* 'room'	p^h*ang* 'bang'
	dal 'daughter'	*tal* 'moon'	t^h*al* 'mask'
	gæta 'to break'	*kæta* 'to fold up'	k^h*æta* 'to dig'
	[Voice]		[Spread Glottis]

The consonants in the righthand column are *aspirated*, and they therefore receive a feature **[Spread Glottis]**. (Notice that the initial plosives in English words such as *poet, taste* and *castle* are also aspirated; but there are no minimal pairs between aspirated and non-aspirated plosives in English.)

Vocalic features

We can now turn to vowels. Just like we can distinguish between two major groups of consonantal features — manner and place of articulation —, we can also also distinguish two types of vocalic features: those of aperture and those of place.

Aperture

Aperture features describe the degree of opening of the jaw. Although there have been claims of languages showing four or even more degrees of opening, it is commonly assumed that most languages have at most three such degrees, and these are usually formalised with two features, **[High]** and **[Low]**. We then get the following tripartition (the following might describe the five vowel system of Modern Greek or many other languages around the world): **Aperture features**

(20) a. High vowels like [i] and [u] have the feature [High]
b. Mid vowels like [e] and [o] do not have an aperture feature
c. Low vowels like [a] have the feature [Low]

The idea is that no vowel can have the features [High] and [Low] at the same time, since these give opposite instructions to the tongue. In this way, we can thus derive a three-way distinction with two features.

Another feature which is sometimes also considered 'aperture' is **[ATR]**, for *Advanced Tongue Root*, describing a movement for the back of the tongue. This feature is very often used to describe the difference between e.g. [e] and [ɛ] or between [o] and [ɔ] in Bantu languages (the first of these pairs have [ATR], but the second does not). The similar vowels in English are also sometimes described in this way, although it is more difficult to detect actual movement of the tongue root on these vowels. **Advanced Tongue Root**

Vocalic place

Like for consonants, the most important articulators for vowels are the lips and the front and the back of the tongue. I adopt a tradition in which the places of articulation are expressed by the same features (with the same names) as those of consonants. You should be aware that there is no absolute consensus on this, and there are many other names for these features, and I will mention some of them here as well; however, in the other chapters of this book, I will only use the feature names mentioned first here.

The three main places of articulation for vowels, then are [Labial], [Coronal] and [Dorsal]. I will discuss these in turn.

Labial vowels are also called *round* (in which the feature might als be [Round]). **Labial vowels** One difference with consonants is that it is not unusual for vowels to carry more than one feature. In particular, lip rounding can go together with some movement of the tongue. For acoustic reasons, this usually is the back of the tong: raising that back has a similar effect on the sound of the vowel as rounding the lips. Examples of [labial, velar] vowels are [u] and [o] of which in particular the first ([Labial, Dorsal, High]) is extremely common in languages of the world.

The combination [Labial, Coronal] is much rarer. French [y] (the vowel in *tu* 'you') is an example, but English does not does have this vowel — although some modern varieties of British English come rather close to it in their pronunciation of [u] (as in *you* [jy]). The reason why this type of sound is presumably rarer is that raising the front part of the tongue has an acoustic effect which is very different from that of rounding the lips. While tongue and lips thus reinforce each other when pronouncing [u], they give partially conflicting signals when pronouncing [y]. Under these circumstances, the former vowel has a better chance of survival than the latter.

front vowels

A vowel which has only [Coronal] are the so-called *front vowels* ([Front]). Examples of these are [i] ([High, Coronal]), [e] ([Coronal] without aperture specification) and [æ] ([Coronal, Low]). When they are not also labial, these sounds tend to be produced with spread lips.

Back vowels

Back vowels, on the other hand, have the feature [Dorsal], which tends to cooccur with [Labial] for the acoustic reasons just explained. Most English back vowels are also labial (we already mentioned [u, o]). Only the low vowel [a] is not round — which may have something to do with the fact that low vowels are pronounced with a rather wide opening of the jaw, which makes it difficult to round the lips at the same time. [a] can thus be specified as [Vocalic, Low, Dorsal].

placeless vowels

schwa

Vowels can also be placeless, i.e. missing all three of {[Labial], [Coronal], [Dorsal]}. The most well-known of these *placeless vowels* is *schwa*, the vowel in the second syllable of *better*. Its IPA symbol is [ə]. There is also (at least) one high placeless vowel, [ı], and a low placeless vowel, [ʌ].

Vowels are sometimes also differentiated according to the pitch with which they are pronounced. For instance, Yoruba (Niger-Congo, Nigeria/Benin/Togo) has a minimal pair such as the following:

(21) *ọ̀wọ̀* 'honour' - *ọ̀wọ́* 'group'

tone

The grave accent (*à*) denotes a relatively low *tone*, while the acute accent denotes a high tone. The word for honour and group are thus differentiated only by these tones, which can be modeled by the features **[High]** or [H] and **[Low] or [L]**. Tone languages will play an important role in the discussion in Chapter 4.

The final vocalic feature I want to mention is again shared with consonants: [Nasal]. For vowels, it surfaces in languages like Saraiki (Indo-European; Pakistan, India, Afghanistan), which contrast nasal from oral vowels, as the following (near-)minimal pair shows:

(22) a. [aːʋi] 'forge for preparing bricks'
b. [ãːʋ̃ĩː] 'you should come/do come'

Nasalisation is indicated by a tilde on top of the vowel. In (22b),also the consonant [ʋ] is nasalised. Languages which have nasal vowels also always have nasal consonants. Furthermore, the nasality of vowels is often shared with adjacent consonants.

Monovalence and markedness

privative

Finally, a word of warning. There are two interpretations of phonological features. Here I have presented an interpretation of features as *privative*: high vowels have the feature [High], but non-high vowels do not. When you go on to study phonology articles, you will discover that some authors have a different representation of features. For them, all segments possess all features, but with different *values*, usually denoted by + and − signs. Under such an interpretation, high vowels have a feature [+high] and non-high vowels have a feature [−high]. This interpretation of features is called *binary* for obvious reasons.

values

binary

To a large extent, these two different notations are equivalent, but there are some differences as well. One reason why I, like many modern phonologists, prefer the privative notation is that it expresses more clearly that the two 'values' of a feature are not symmetric. A non-nasal vowel is in several ways simpler than a nasal one. All languages have oral vowels, but not many have no oral ones. Furthermore, even if a language does allow for nasal vowels, their occurrence is much more heavily restricted than that of oral vowels. For instance, they can only occur next to nasal consonants, as in the Saraiki example in (22b). Such properties together are summarised by phonologists in saying that nasal vowels are more *marked* than oral vowels.

marked

Such asymmetries are better expressed in a theory in which nasal vowels have an extra feature which is absent in oral vowels than in a theory in which they both have the same feature, only differing in their vowels. The former are literally more complex in our representations, which correspond to their being treated as more 'difficult' in human language. In a theory with binary features, there is no formal difference: a nasal vowel is [+nasal] and an oral vowel is [−nasal], which means that there representations are equally complex. (See section 4.5 for more discussion of the notion of 'markedness'.)

2.4 An alternative: Element Theory

The aim of this book is to introduce you to the most important ideas in the mainstream of present-day theoretical phonology. As in any field of research, there is however no absolute uniformity on every point; to the contrary, many points of the theory are still very much debated, and every assumption is questioned from time to time — as it should be.

It is therefore useful to consider an alternative approach to some of the assumptions underlying the mainstream model presented so far. This is so-called *Element Theory*, a theory which assumes primitive elements that are similar to, but not exactly the same as the features we have seen. They are also not called 'phonological features', but phonological *elements* instead.

Element Theory

elements

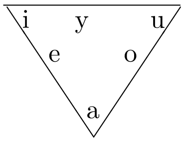
An important difference between elements and features is that the latter cannot be pronounced. A feature such as [Labial] does not correspond to any single acoustic or articulatory event in isolation; we always need to build a segment with many other features in order to produce it. Elements on the other hand, can be independently pronounced.

Most vowels in the world's languages, for instance, will consist of the elements |A|, |I| and |U| in some constellation. (We will concentrate in this section on the phonology of vowels, because this is what a large amount of work within this framework has been devoted to; this is not to say that the theory has not been successfully applied to the analysis of consonants as well, however.) Elements are usually spelled with a capital letter, and they are placed between | | brackets in order to distinguish them from features and phonological or phonetic strings of segments.

Each of these elements can be pronounced in isolation:

(23) a. | A | is pronounced as [a]
b. | I | is pronounced as [i]
c. | U | is pronounced as [u]

Other vowels can be analysed as a combination of one or more of these basic, most primitive elements — it is of course not a coincidence that the elemental vowels in (23) correspond to the three angles of the vowel triangle:

(24)


vowel triangle

The *vowel triangle* is a graphic representation of a vowel set in languages of the world. The triangle can be seen as an abstract graphic representation of the mouth, with the lips on the right-hand side and the back of the mouth on the left-hand side. (The vowel triangle also represents acoustic properties of these vowels; see the section on Further Reading on page 45.)

Vowel systems tend to have this triangular shape. If a language has only three vowels (such as Classical Arabic), these will be typically [i, u, a]. If it has five, the set will be [i, u, a, e, o], and in this way the vowel triangle gets more densely filled the more vowels we have.

cardinal vowels

This generalization on vowel systems is represented rather nicely in an element system. The three quasi-universal vowels are the corners of the triangle, and other vowels consist of combinations of these '*cardinal vowels*'. For instance, in a typical five vowel system we will have the following combinations:

(25) a. The combination | A | • | I | (or | I | • | A |) is pronounced as [e]
b. The combination | A | • | U | (or | U | • | A |) is pronounced as [o]

The order in which we present the combinations of features is of course irrelevant: it does not matter whether we write | X | • | Y | or | Y | • | X |, since both refer to the same phonological representation. (In mathematical terms, the operation • is commutative.)

The combination | U | • | I | will be typically pronounced as [y], and the mid front rounded vowel [ø] would consist of the combination | U | • | I | • | A |. The system is built on the typological observation that three vowel systems usually occupy the three corners of the vowel triangle. All other vowels are typologically more 'marked': they exist in fewer languages.

This observation cannot be expressed directly in a theory which uses features. It is not part of the formal system of ordinary feature theory that the [High,Coronal] vowels ([i]) are much more frequent than [Coronal,Low] vowels ([æ]) in languages of the world. This is something which just needs to be stipulated, or derived from the phonetic difficulty to pronounce the latter type of vowel. Within Element Theory, the issue becomes clear from just looking at the different representations of the two types of vowels, and the theory therefore seems more restrictive.

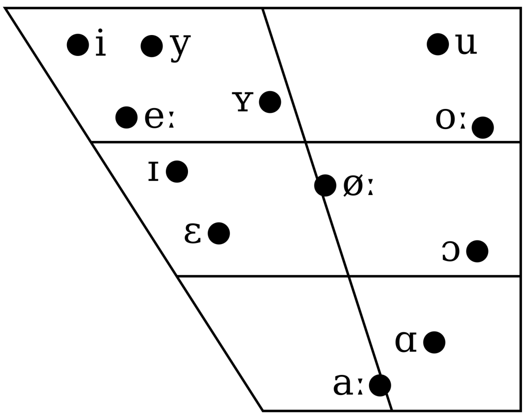
Notice, on the other hand, that Element Theory does not provide us with an answer to the question why [e] and [o] are typologically much more frequent than [y], and why the latter often behaves as a less frequent vowel in languages which have it, so that not every typological question is answered.

A popular metaphor of elements in the phonological literature is that of colours. We only need the three primary colours *red, yellow*, and *blue* to produce all other colours by mixing them in the appropriate quantities. In the same way, we can derive (almost) all vowels from the three elements.

The vowel system of Dutch

At first sight, Element Theory might seem too restrictive in an obvious way. Given three elements, we can only derive six vowels — the ones we have just mentioned. But many languages have many more vowels. For instance, Dutch has 13 vowels (not counting three diphthongs and some vowels which only occur in loanwords):

(26) i, y, u, e, ø, o, a, ɪ, ɛ, œ, ɔ, ɑ, ə



(27)

(The schwa is missing from the vowel triangle, but we will return to it shortly.) We cannot go into all the complications of the Dutch vowel system, but we can illustrate some of the strategies within Element Theory. In the first place, in our colour analogy, we mentioned that we can mix colours in the appropriate quantities.

For phonological elements, we can express this by introducing the notion *headedness*. When we combine two linguistic elements — two words in a syntactic phrase, two morphemes in a word, two syllables in a stressed unit, etc. — we can always give a special status to one of them: this one is the *head*. **headedness**

We can now extend this idea to phonological elements: if we combine them, we can assign the head status to one of the two. This doubles our representational possibilities. If we have a combination of two elements |X|•|Y|, we can distinguish between $|\underline{X}| \bullet |Y|$ and $|X| \bullet |\underline{Y}|$, where the underlining denotes the head of the segment. Thus we get the following distinctions within the realm of mid vowels:

(28)
a. The combination $|A| \bullet |\underline{I}|$ is pronounced as [e]
b. The combination $|\underline{A}| \bullet |I|$ is pronounced as [ɛ]
c. The combination $|\underline{A}| \bullet |U|$ is pronounced as [ɔ]
d. The combination $|A| \bullet |\underline{U}|$ is pronounced as [o]

[ɛ] and [ɔ] are lower than their counterparts, therefore they are more $|A|$-like, and they have this element as their head. The head of a segment thus is the one which has the strongest influence on the phonetic result.

The result obtained so far works very nicely for the many languages which have a seven-vowel system: they usually have the four vowels in (28), next to the three primary vowels of course. We are then still assuming a system in which $|I|$ and $|U|$ cannot be easily combined; notice, by the way that Swedish is a language which distinguishes between two front rounded vowels, and thus features a headedness distinction in $|I| \bullet |U|$ combinations.

However, the system built up so far is certainly not sufficient for Dutch, since this language still has almost twice as many as seven vowels. A solution here comes from the study of one of these, the schwa ([ə]). This vowel is hard to describe in terms of the elements we have seen so far: it is the central vowel, right in the center of the vowel triangle and from an articulatory point of view it is 'targetless': it does not seem to involve the active use of any specific supralaryngeal articulatory organ.

Although we should be very careful in introducing new phonological elements — because that would run against the spirit of the program, which requires us to be as restrictive as possible — the special behaviour of schwa seems to warrant the introduction of a new element, $|@|$ (the @ sign was sometimes used in the past as an alternative to 'ə' in cases where the latter was not available, e.g. when using a computer that did not yet have the option of representing phonetic letters; basically all modern computers do have that possibility, as the IPA is part of the so-called Unicode character set, but the tradition still remains).

$|@|$ is special because it is targetless, and therefore it does not have any effect on the realisation of the vowel.

(29)
a. $|\underline{A}| \bullet |@| = |A|$
b. $|\underline{I}| \bullet |@| = |I|$
c. $|\underline{I}| \bullet |A| \bullet |@| = |I| \bullet |A|$
d. ...

In this sense, it behaves like 1 in multiplication or 0 in addition, which also do not change the end result, the technical term in mathematics is *identity element*):

identity element

(30)

$1 \times 1 = 1$	$1 + 0 = 1$
$2 \times 1 = 2$	$2 + 0 = 2$
...	...
$n \times 1 = n$	$n + 0 = n$

Addition of the schwa in this way at first sight does not make our system much more powerful: it behaves as the identity element, so combining it with existing elements does not give new results. But there is one escape hatch: it may be possible to extend the notion of headedness also to structures with a schwa. If schwa is the head of a combination, it *does* have influence on the interpretation: it *centralizes* it, it draws it into the centre of the vowel triangle. This seems to give a proper description of the difference between e.g. [i] and [ɪ] or [a] and [ɑ] in Dutch.

centralizes

We can now give the following matrix of possibilities:

(31)

|A|	[a]	|I|	[i]	|U|	[u]
|A|•|$\underline{@}$|	[ɑ]	|I|•|$\underline{@}$|	[ɪ]	|U|•|$\underline{@}$|	✘
|A|•|$\underline{I}$|	[e]	|A|•|$\underline{U}$|	[o]	|I|•|$\underline{U}$|	[y]
|$\underline{A}$|•|I|	[ɛ]	|$\underline{A}$|•|U|	[ɔ]	|$\underline{I}$|•|U|	✘
|A|•|I|•|$\underline{@}$|	✘	|A|•|U|•|$\underline{@}$|	✘	|I|•|U|•|$\underline{@}$|	✘
|I|•|$\underline{U}$|•|A|	[ø]	|$\underline{I}$|•|U|•|A|	✘	|I|•|U|•|$\underline{A}$|	✘
|I|•|U|•|A|•|$\underline{@}$|	[œ]	|@|	ə		

The crosses ✘ in this table denote segments which could be produced given the combinatory rules but for which we do not have evidence for in the Dutch system (note that we have rather arbitrarily assigned the head status in some combinations of elements). We could try to find some reason for why certain combinations are lacking — maybe there is some reason why |U| always needs to be a head when it combines with |I|, or why |@| does not seem to be the head in complex expressions (except, strikingly, the most complex one of all: |I|•|U|•|A|•|$\underline{@}$|).

Be this as it may, we still predict 20 possibilities by the formal system alone, while we find only 13. However, if we compare this to a standard feature theory of the same inventory, this fares relatively well. Within such an analysis, we would need at least the following five features for Dutch: [Dorsal], [Labial], [ATR], [High], [Low]. Since all these features are binary, we have $2^5 = 32$ logical possibilities. The element model thus gives a tighter fit to the data.

Vowel reduction

In many languages of the world, there is an interesting difference between stressed and unstressed positions of the word: in stressed positions we usually find a larger number of phonological contrast, which is *reduced* in unstressed position. In Belorussian (Indo-European, Belorussia), we find {i, u, e, o, a} in stressed syllables, but only {i, u, a} in unstressed position. This distributional preference is also responsible for alternations: if an /o/ or /e/ ends up in an unstressed position, it will be reduced to [a]:

reduced

(32)

stressed	*unstressed*
nóɣi 'legs'	*naɣá* 'leg'
kól 'pole NOM'	*kalá* 'pole GEN'
v^j ósnɨ 'spring GEN'	*v^j asná* 'spring NOM'
ʃépt 'whisper'	*ʃaptáts^j* 'to whisper'
kl^j ej 'glue'	*kl^j ajónka* 'oil-cloth'

centrifugal

centripetal vowel reduction

This type of reduction is sometimes called '*centrifugal*': the vowels move to the corners of the vowel triangle when they do not carry stress. It can be opposed to *centripetal vowel reduction*, in which all the vowels seem to move to the center of the vowel triangle in exactly the same types of positions. An example is (informal) Dutch, in which any kind of unstressed vowel can be reduced to schwa — a further difference with Belorussian is that the process is optional in Dutch, but that is immaterial to our present discussion.

(33)

formal	*informal*
fonologie 'phonology' [fònoloɣí]	[fònələɣí]
minuut 'minute' [minýt]	[mənyt]
kantoor 'office' [kɑntɔ́ːr]	[kətɔ́ːr]

The distinction between 'centrifugal' and 'centripetal' is not very strong in natural language. Some languages show both processes. For instance, Catalan /e, a/ reduce centripetally to [ə], whereas /ɔ, o/ reduce centrifugally to [u] (and /i, u/ do not reduce at all):

(34)

stressed	*unstressed*
sérp 'snake'	*sərpəntí* 'winding'
pél 'hair'	*pəlút* 'hairy'
gát 'cat'	*gətét* 'kitten'
λúm 'light'	*λuminós* 'luminous'
gós 'dog'	*gusét* 'puppy'
pɔ́rt 'port'	*purtuári* 'of the port'

How can we understand these processes, and the fact that together they seem to give a fairly complete catalogue of vowel reduction phenomena in languages of the world? Within Element Theory it is very easy to see what is going on. All reduction processes are an instance of the following:

(35) An unstressed vowel is not allowed to carry more than one element; delete elements if necessary.

prominent

What is more, (35) itself can be understood as a sensible restriction on phonological structures. In an intuitive sense, unstressed syllables are less *prominent* than stressed ones: they are not so loud and in several other ways less important. From perception experiments, we know that listeners pay less attention to them. This will be formalized — in chapter 7 — as another headedness relation: stressed syllables are heads, unstressed syllables are not heads. Some authors equal vowel reduction to 'information loss': unstressed positions can carry less information than other positions. It is not strange that information may get lost in positions that are not so prominent.

Centrifugal reduction now leads to one of the primary elements |I, U, A|; centripetal reduction leads to |@|. Notice that we still need to build an asymmetry into our system: /i, a, u/ may reduce in some languages to [ə], but /ə/ will never reduce to any other vowel, not even a simple one. This should of course have some relation to the fact that [ə] is the 'identity element'. The fact that |F|=|F|•|@| means that in some sense |@| is part of all vowels, but inversely, none of the primary vowel elements are part of [ə]. Also in this sense, then, |@| behaves like the mathematical zero. The fact that schwa is the targetless vowel makes it the one which carries the smallest amount of information — hardly any at all. That may make it the best

Feature theory has many other virtues, and is therefore an important topic for any student of phonology. We will continue to use this theory, rather than Element Theory, as the background for the following classes. It is important, however, to realize, that fruitful alternatives exist to many of our key assumptions.

2.5 Exercises

1. In the 19th Century, the French music teacher François Sudre invented a language, Solresol, which had only seven distinct sounds: *do, re, mi, fa, sol, la, si*. Such a language has obviously many advantages, for instance because you can also play words on a musical instrument. Still, Solresol was not very succesful and no language has this minimal kind of inventory. Mention some disadvantages to organizing sounds in this way.
2. The vowel [i] may have a slighly different set of features in a language which only has the three vowels {i, a, u } than in a language which has the five vowels {i, e, a, o, u }. Explain.
3. Consider the vowels of Turkish: { i, o, a, ø, u, y, e, ı }. Display these vowels in a table which shows how you can describe the set in terms of features. You may assume that ırepresents a high schwa-like vowel.
4. Now do the same exercise in terms of elements.
5. Consider once more the speech errors in (17). Describe exactly which features have been moved from one consonant to another.
6. For each of the following segments, give a full feature specification: [f, h, ʝ, ø, ŋ, p, ʋ, ɣ, ə, e]. (Look the sounds up in an IPA table if you do not know them.)
7. For each of the following feature combinations, give a corresponding IPA symbol.
   (Ła) [Consonant, Sonorant, Labial, Nasal]
   (Łb) [Consonant, Sonorant, Coronal, Lateral]
   (Łc) [Consonant, Fricative, (Coronal), Dental, Voice]
   (Łd) [Consonant, Stop, (Coronal), Dental, Voice]
   (Łe) [Vowel, Coronal, High]
   (Łf) [Vowel]
8. For each of the following sets of sounds, describe what feature(s) they have in common:

   (Ła) [b, d, v, β, z, g]

(Łb) [m, n, ã, ŋ, ɱ]
(Łc) [o, u, y, w, ʏ]
(Łd) [b, m, f, p, ɱ]
(Łe) [m, r, l, ɾ, n, j]
(Łf) [i, a, u, y, e]

9. The prefix *in-* in English has several different forms. Before a plosive it takes the following shapes:

   (Ła) *i*[n]*active*
   (Łb) *i*[n]*decisive*
   (Łc) *i*[n]*secure*
   (Łd) *i*[m]*popular*
   (Łe) *i*[m]*balance*
   (Łf) *i*[ŋ]*coherent*

   Describe what happens in terms of features.
10. An interesting way of studying the articulation of speech sounds is the *ultrasound* technique, which is relatively non-intrusive and allows us to see the movements of the tongue. Speech scientists at the University of Glasgow set up a YouTube Channel which represents many different sounds. You might be able to find more clips elsewhere. Study a clip for the uvular, bilabial and palatal nasal. Make a still — a non-moving picture — at the moment in which each of these is pronounced, and show how you can indeed see the different places of articulation as described in this chapter.

ultrasound

11. Take the consonant inventory of a random language of your choice that has not been described in this chapter (for instance, your native language, or some other language you might be interested in). Describe this inventory in terms of the features we have seen so far. Do you find any segments which need features that we have not yet introduced?
12. Japanese speakers who learn French sometimes replace the [y] sound of the language, by the sequence [ju]. Explain why this might be the case.
13. In the text (p. 2.4) it is suggested that the Dutch vowel system would be more difficult to describe with binary features. Work out such an analysis, and show how it is more complicated than the one in terms of elements.
14. Give an analysis of Belorussian vowel reduction in terms of binary features rather than elements.
15. Data in some typological databases, such as WALS, suggest that (almost) all languages have voiceless plosives, whereas a much smaller number have voiced plosives. Discuss how this gives evidence for privative features.
16. What is the simplest vowel according to feature theory? What about Element Theory? What kinds of evidence would you explore to see whether such hypotheses are justified?
17. Lakhota (Siouan, North America) has five oral vowels ([i, e, a, o, u]) and three nasal vowels ([ĩ, ã, ũ]). Explain why we would not expect a language which would have a distribution the other way around (five nasal vowels and three oral vowels) given markedness theory.

18. The online edition of the *WALS* (World Atlas of Language Structures) has, among other things, a chapter on front rounded vowels, written by Ian Maddieson. On the map in this chapter, you can see that front rounded vowels are very rare in languages of the world. Further, Maddieson distinguishes between two types of front rounded vowels, mid and low. Element Theory and Feature Theory make slightly different predictions about which of these two should be more frequent. Do these data provide any evidence for either of these theories?
19. The *UPSID* database (UCLA Phonological Segment Inventory Database) contains overviews of the segments of several hundreds of languages. Find in this database, languages which have:

    (Ła) [p]
    (Łb) [y]
    (Łc) [d]
    (Łd) [ŋ]
    (Łe) [l]

Sources and further reading

Section 2.1 Many examples about the relevance of orthography to phonology are taken from Coulmas (2003). More discussion about the difference between consonants and vowels from a psycholinguistic point of view can be found in Bonatti et al. (2005). The brain scanning data about the same topic is found in Carreiras and Price (2008). Several scholars have been important for introducing the relevance of Semitic (and other) templates into phonological theory. Very important was McCarthy (1979)'s dissertation; Bat-El (2011) gives a good overview of the literature on several Semitic languages, such as Arabic and Hebrew.

Section 2.2 The role of features in acquisition has been argued for by Levelt and van Oostendorp (2007). The examples of speech errors are from a famous classical article by Fromkin (1973). Lahiri and Reetz (2010) give an excellent overview of psycholinguistic evidence for phonological features. An interesting overview and development of the concept of contrast can be found in Dresher (2009).

Section 2.3 The Saraiki data are from Syed (2012). Several phonologists have expressed doubts in recent years as to the universality of the feature set; prominent among those is Mielke (2004). (Most of these scholars would agree that the features presented here give a good approximation for most phenomena in most languages.)

Section 2.4 Element Theory has been worked out in particular in the framework of Government Phonology (Kaye et al., 1985, 1990; Harris, 1994; Harris and Lindsey, 1995); related views were also found in the framework of Dependency Phonology (Anderson and Ewen, 1987). An excellent introduction into this subject matter is provided by Backley (2011). The Belorussian data are from (Crosswhite, 2001; Harris, 2005). The latter author is also responsible for the suggestion that vowel reduction is information loss.

Finding patterns

3.1 Two aspects of phonology

Theories of grammatical form typically consist of two parts: a theory of *representation* and a theory of *computation*. The former is the one we have been mostly engaged with so far: it studies how an individual linguistic representation is structured: a word consists of segments that in turn are structured in terms of features. THe theory of computation complements this with ideas about how individual representations are related to each other.

representation
computation

This is relevant in particular in phonology if we consider those two forms to be related or even the same at some level. Consider the case of the plural in English. Sometimes the plural is formed by an [s] (*cats*), sometimes by a [z] (*dogs*) and sometimes by [əz] (*fishes*). At some level we can see these as instances of the same thing: they all express plurality, and there representations have a lot of things in common (the features of a coronal fricative). There are also differences, and these differences seem to be motivated by the different phonological context in which they occur: the plural is voiceless if the immediately preceding obstruent is and voiced otherwise; we see a schwa in the plural only if the stem ends in a coronal fricative.

The theory of representations can help us understand this, but only partially. English does not have sequences of obstruents with different voicing, or sequences of two fricatives with the same place in a row, so it makes sense that this thing which at some level is the same — 'the plural morpheme' — takes the different shapes it has. But we still want to also be able to say that these forms are indeed instances of 'the same' thing, because they are related to each other, and we have reason to believe that relation is also stored in the lexicon. That is what the theory of computation is about.

There are two kinds of theories about this. One theory holds that [s], [z] and [əz], are all stored in the lexicon and so is some kind of link between them; another theory says that only one of these is stored (the *underlying form*) and the others are derived from it by *derivation*: some regular phonological process turns the underlying form into something else in some context.

underlying form
derivation

The consensus among phonologists probably is that for certain cases we need one and for others we need the other theory. Sometimes, one form reacting to different contexts can show behaviour that seems no longer regular to assume that it plays any role in the internalized grammar of the native speaker. There is a correspondence in meaning between *tooth* and *dentist*, and there is even some correspondence in form because the two historically derive from the same, Indo-European, root of some thousands of years ago, but most phonologists would be unwilling to consider that when e.g. saying *tooth*, we start from the 'underlying form' /dens/ and then adapt it to the phonological environment which results in [tuθ]. An even more extreme example would be (the beginning of) *gynaecologist* and *queen*: there is a meaning relation (woman) and some similarity in form (a front vowel, an n), all of this due to a very distant etymological relationship. But it is not clear that these similarities play a role in the contemporary language system or that these forms are in any way very closely related. We therefore tend to consider this relation as accidental.

productive

On the other hand, we have evidence that not every individual form of a morpheme is just stored like that. Certain morphemes are *productive*, and the plural affix is an example of that. We can add it to any newly formed noun, and it will always show up in the right shape. A famous example is the so-called *Wug test*: you show a person a picture of some fantasy animal and tell them it is called a *wug*; then you show them two of those animals and ask what they would be called. Every native speaker of English would tell you they are two *wugs*, with the fricative being voiced. This means that speakers of English have knowledge of which shape of the plural goes where, and this is nicely captured by the model of the underlying form and the derived forms.

allomorphs

We thus conclude that probably both are necessary: sometimes different *allomorphs* (i.e. shapes of the same morpheme) are stored, denoting the same thing; at other times we store only one underlying form, and derive other forms by phonological computation, adapting sounds to their environment.

In this chapter and in chapter 6 we will be concerned mostly with this issue. In this chapter, the main issue is to find underlying representations; in chapter 6 we then study the way in which other forms are derived from the underlying representation.

3.2 Underlying representations

First step

final devoicing

Let us first concentrate on one fairly simple process, *final devoicing* in German, illustrated in (36) (we have seen a similar example for Turkish in (11) in Chapter 2):

(36) a. *Hund* 'dog (sg.)' [hunt], *Hunde* 'dogs (pl.)' [hundə]

The first observation we have to make is that we know that [ə] is a plural suffix in German. We also find it for instance in *Schuh* 'shoe (sg.)' [ʃu] - *Schuhe* 'shoes (pl.)' [ʃuə]. We thus have a situation where we have to assume that the forms [hunt] and [hund] are derived from the same underlying form.

As a rule of thumb, we can safely assume, at least at the current level of discussion, that the underlying form always is a surface form somewhere, so that we have two possibilities: either /hund/ is the underlying form or /hunt/ is. The other form will then be derived. It is useful to make a notational point here: we write underlying forms between dashes //, whereas output forms are written in square brackets []. A further notational point is that we write a derivational relation between two forms with an arrow, pointing from the underlying form to the phonetic form, as follows:

(37) /α/ → [β]

So which of the two forms /hund/ or /hunt/ is the underlying representation for the German word for 'dog'? In principle, both are possible and the isolated fact in (36) does not provide us enough information to decide. We have to consider the overall system of German to make a choice between the two possible relations between underlying and surface forms 38a and 38b (such relations are often called derivations):

(38) a. /hund/ → [hunt] (at the end of the syllable)
b. /hunt/ → [hund] (before ə)

How do we choose? One observation we can make when we study more facts from the language is that there actually are no German words at all ending in a [d] (or any other voiced obstruent, for that matter). Loanwords into German will also adapt to this generalization. This is captured by the derivation in (38a): even if a word would hypothetically have an /d/ in the lexicon, it will never make it to the surface as a [d], but it will always become a voiceless consonant, following the change prescribed in (38a). Derivation (38b), however, has nothing to say about this fact; it has to be stated elsewhere.

Also, (38b) cannot easily be extended to other German facts. For instance, there are many German words in which the /t/ does not change before a schwa. For instance, the adjective for 'colourful' is *bunt*. When we inflect this adjective, it can get a schwa, but the /t/ remains unchanged: *bun*[tə]. This would mean that (38b) has many exceptions, but (38a) has none.

Although descriptively for the one individual fact in 36 both analyses are equivalent, we aim to choose the analysis that is easiest to generalize to other facts and which raises the smallest number of exceptions. Since it is easy to generalize (38a) (viz. to (39)), but not (38b)), we assume that the former holds, and hence /hund/ is the underlying form:

(39) /. . . [voice]/ → [. . . ∅] (at the end of the syllable)

The condition 'at the end of the syllable' is also a familiar one and has a property of phonetic naturalness: at the end of syllables (or words) it is more difficult to produce voiced obstruents and to hear the difference between voiced and voiceless obstruents. Such obstruents are thus under constant pressure to neutralize, and it is more likely that languages develop the process in (38b) than that in (38a).

We thus have the following rough recipe for determining the underlying form of a morpheme in a data set of *polymorphemic forms* (forms consisting of more than one morpheme): **polymorphemic forms**

(40)
1. List all forms of the morpheme in the data set
2. For each of these, analyse what should change to it in which contexts, in order to derive all the other shapes of the morpheme (i.e. list possible derivations).
3. Of these analyses, choose the one that is (a) simplest, (b) easiest to generalize to (potential) other cases, (c) most natural in phonological terms, (d) has the smallest number of exceptions.

4. The form corresponding to the analysis in step 3 is the underlying form.

It may seem that there is something inherently subjective in step 3, and even though we will not attempt to set up a precise, formalized, 'objective' set of criteria to decide what is the 'simplest' analysis, and indeed in the literature one might find (complicated) cases where opinions diverge. However, in most cases phonologists' opinions converge as it is indeed easy to see what is simple: a simpler description will tend to be shorter for instance.

Complementary distribution

alternation

complementary distribution

allophones

phoneme

The above example of finding an underlying form in German is based on polymorphemic forms. In such cases we can see an *alternation*: the same morpheme shows up in different shapes in different environments. However, there is a potential other source for our knowledge of phonological computation: so-called *complementary distribution*. Languages may sometimes dispose of two different sounds, call them A and B, where A only occurs in one set of environments and B only in another ('complementary') set of environments. In that case we tend to assume that A and B are different instantiations of the same sound: in traditional phonological parlance it is said that they are *allophones* of the same *phoneme*.

It is methodologically more difficult to find patterns of complementary distribution than it is to find active patterns of adaptation of morphemes: in order to be really sure that words are in complementary distribution one needs to check the whole lexicon. However, whenever you find a minimal pair, you can be sure that the sounds are *not* in complementary distribution.

In the Southern dialect of Kikongo (Bantu, Congo), we find among others the following words.

(41) *tobola* 'to bore a hole', *čina* 'to cut', *kesoka* 'to be cut', *nkoši* 'lion', *zenga* 'to cut', *žima* 'to stretch', *kasu* 'emaciation', *čiba* 'banana', *nselele* 'termite', *lolonži* 'to wash', *zevo* 'then', *ažimola* 'alms', *nzwetu* 'our house', *kunezulu* 'to heaven', *tanu* 'five'

If we look closely at these forms, we discover that the coronal sounds [t,s,z] and their palatal counterparts [č, š,ž] occur in complementary distribution. The latter occur before the high front vowel [i] and the former elsewhere.

We can thus assume that [t] and [č], [s] and [š], and [z] and [ž] are in complementary distribution. As before we now have to determine which is the underlying form, and as before we choose the form that would give us the simplest description of the language in question. Compare the following two (just for [s]-[š] for convenience)

(42) if /s/ is underlying:
/s/ → [š] before [i]

(43) if /š/ is underlying:
/š/ → [s] before [e, o, u, a]

Clearly (42) is simpler than (43), already in the sense that it is shorter. The latter rule also cannot be further simplified, as the vowels [e, o, u, a] have nothing in common except that they are not [i]. To people who are familiar with linguistic typology, it is also the case that the process in (42) is found in many languages of the world, and known as *palatalisation*. The process in (43) would be depalatalisation; it is also known, but much rarer, and as far as we know it is not found in this complicated disjunctive context. Both processes are to some extent phonetically natural since the (palatal) consonants are produced with the tongue in the front part of the mouth, just like the front vowels. Both processes also can obviously be generalized to all coronal obstruents (Southern Kikongo does not have a voiced coronal plosive).

palatalisation

If we accept the complementary distribution as sufficient evidence, we can thus assume that the non-palatal obstruents are underlying, and the palatals are derived. As a side note, it should be mentioned that not all phonologists accept this type of evidence for assuming an active synchronic phonology. An alternative explanation for the distribution is that (42) has been a historic process, affecting items in the lexicon at some point, which since then have been stored in their current form — some with a palatal and others with a plain coronal. In this case, the phonological analysis would only apply to the historic stage when the change actually happened.

Sometimes we find patterns of complementary distribution that seem more accidental — they have no origin in history and they are difficult to understand from a phonetic point of view. A classic case in English is the distribution of [h] and [ŋ]: the former occurs in the onsets of stressed syllables, the latter elsewhere:

(44)

help	*ŋelp
hero	*ŋero
*siher	siŋer
*sih	siŋ

[h] and [ŋ] have no common historic origin and phonetically they are quite different — a phonological process would have to posit quite a lot of changes to go from one to the other. It is therefore usually assumed that there this complementary distribution does not have to be accounted for by phonology (although we do want to explain why both sounds are restricted in their respective distributions).

Which morpheme is changing?

In many cases, the phonological analysis will have to be performed in tandem with a morphological analysis in order to succesfully decide about the phonological shape. Take for example the following data from Farsi (Indo-European, Iran):

(45)

	SINGULAR	PLURAL	gloss
a.	zæn	zænan	woman
b.	læb	læban	lip
c.	mæleke	mælekean	queen
d.	valede	valedean	mother
e.	setare	setaregan	star
f.	bænde	bændegan	slave

Suppose we wanted to have the underlying forms for each of these morphemes. For the first four examples, there does not seem to be a problem: in order to make the plural of 'woman' we take the singular form *zæn* and add *an*, so a reasonable first hypothesis is that the word for 'woman' is *zæn* underlyingly and the suffix is *an*. However, something changes in the last two forms of the table: in the plural we there see a *g* which is not there in the singular form of the nouns in question or in the form we have thus for hypothesized for the plural suffix.

We are now faced with three logical possibilities:

(46)
1. The [g] is not part of any morpheme, but it is inserted for some reason in the plural contexts of (45e) and (45f).
2. The /g/ is part of the suffix and it is somehow deleted in the plural forms of *star* and *slave*.
3. The /g/ is part of some stems and it is somehow deleted in the singular form.

All three of those purported analyses correspond to attested phenomena in languages of the world. A possible analysis along the lines of (46-1) would be that the g would be inserted between two vowels. This is typologically not implausible, as languages tend to like regular alternations of vowels and consonants. A similar analysis assuming (46-2) would be that the suffix is underlyingly *gan* and the /g/ is deleted after a stem-final consonant.

Both analyses suffer from the fact that there are stems in (45c) and (45d) that also end in a vowel in the singular but do not show the [g] in the plural. It is not clear what would be the distinction between those forms from the ones in (45e) and (45f) that could explain why insertion or deletion would apply in one set of cases and not in the other.

We are thus left with (46-3). This would mean that the word for 'star' is *setareg* underlyingly and the /g/ would be deleted in the left column, for instance because it occurs at the end of the word. This analysis does not suffer from the problems we mentioned, as there are no words which do end in a *g* in the dataset (or indeed, in the native Farsi lexicon). It thus is the analysis which is simplest and which we therefore prefer.

The Farsi pattern is typologically a little bit unusual: stems are less likely to undergo changes than affixes and unaffixed words very often represent the underlying form. On the other hand, the German pattern of final devoicing shows something similar: also there we can only see the underlying phonological form in words that are morphologically derived. Such patterns thus definitely seem to occur.

Taking into account the whole grammar

Finding which is the underlying form among the set of surface forms can sometimes be a little tricky. Take the following examples from Kimbundu (Bantu, Angola):

(47) ŋg-a-bit-ile 'I passed' ŋg-a-bet-ele 'I beat'
ŋg-a-batu-is-ile 'I caused to cut' ŋg-a-bet-es-ele 'I caused to beat'

(We will analyse this kind of pattern in Chapter 4 as vowel harmony.) The suffixes *ile/ele* expressing past tense, and *is/es*, expressing the causative, have a high vowel when the stem ends in a high vowel, and a non-high vowel otherwise. The issue is now: which of these forms is underlying, and which is derived? Both seem to have an equally simple grammar:

(48) a. /i/ → [e] next to a non-high vowel
b. /e/ → [i] next to a high vowel

Both rules are equivalent in terms of simplicity and both of them also are equally natural from a phonetic perspective. So how can we decide between these two different ways of seeing the grammar?

We may try to find an answer in the structure of the past tense suffix, which is *ile*, and not *ili*. In other words, the second vowel of the suffix seems to contain a mid vowel that does not undergo the rule in (48b). So that would seem an argument against that analysis and in favour of (48a).

On the other hand, however, the stem vowel remains unchanged in the word for 'I passed', in spite of the fact that it is preceded by a non-high vowel [a] in a prefix. So that rather seems an argument against (48a) and in favour of (48b).

Simplicity therefore does not seem a workable criterion in these cases. However, typological considerations play a role, and we know that adaptation of stem segments to affixes is much more unusual than adaptation of segments within affixes. It thus seems justified to believe that the boundary between prefix and stem blocks certain kinds of processes from happening. From other analyses of Bantu languages we know that this is more generally the case: prefixes behave as if they are much further removed from the stem than suffixes.

This then means we do not have to be as worried by the *a-i* sequence as about the *i-e* sequence in [ŋg-a-bit-ile]. We thus prefer to assume that there is lowering going on, and choose (48a) as our eventual analysis.

Tibetan numerals: splitting morphemes

A spectacular case of morpheme division is the following set of data from Tibetan (Tibetic, Tibet), an assignment in many linguistics olympiads, where participants are invited to find the underlying shapes of the different numerals:

(49)

jig	'one'	*jugjig*	'eleven'		
ši	'four'	*jubši*	'fourteen	*šibju*	'forty'
ŋa	'five'	*juŋa*	'fifteen'	*ŋabju*	'fifty'
gu	'nine'	*jurgu*	'nineteen'	*gubju*	'ninety'

The first thing to note is that e.g. *fourteen* is 'ten four' in Tibetan and *forty* is 'four ten'. But now we have to find the underlying forms for these morphemes, among others. A first hypothesis might be that 'four' is *ši*, as this is the form in the word in isolation. That would mean that 'ten' is *jub* in the word for 'fourteen' and *bju* in the word for 'forty'. This kind of disparity is of course logically possible: the English morphs *teen* and *ty* are also different. However if we look at the other forms it turns out that the form for 'ten' in *eleven* would be *jug* and in 'nineteen' it would *jur*; and so forth. That is a lot of variation: the morpheme meaning 'ten' would basically have a different shape in every word.

If we look more closely at the forms for 'forty', 'fifty' and 'ninety', however, we discover that they all contain *bju* as the string of segments meaning 'ten'. Suppose *that* would be the underlying form of this morpheme. That would mean, first that the initial consonant gets lost in the names of numbers under 20: /bjugjig/→[jugjig], or more generally:

(50) $C_1\ C_2 \rightarrow C_2$ at the beginning of a word

But now that we have uncovered this regularity, we can understand where the consonants in the middle of the names of numbers under 20 come from: from the *second* word. We thus get the following underlying representations.

(51)

gjig	'one'	*bju*	'ten'
bši	'four'		
ŋa	'five'		
rgu	'nine'		

Together with the rule in (50) this describes (the relevant part of) the numeral system in Tibetan. It is not necessary to posit an extra consonant in the word for 'five', since that always shows up in the same way, simply because it starts with only one consonant and this is not subjected to (50).

Cases where the underlying form never surfaces

We can thus extend our algorithm for finding the underlying form as follows:

(52)
1. List all possible forms of all morphemes in the data set
2. For each of these, analyse what should change to it in which contexts to derive all the other shapes of the morpheme
3. Of these analyses, choose the one that (a) gives the simplest overall set of processes for the language in question, (b) easiest to generalize to (potential) other cases, (c) most natural in phonological terms
4. The form corresponding to the analysis in step **??** is the underlying form.

So far, we have assumed that the underlying form is one of the forms that is actually occurring on the surface, but because a given process may touch a form in more than one position, it is possible that the underlying form never surfaces as such. Even though this is a fairly rare occurrence, the Tibetan examples just discussed already give some indication how this might happen.

The following example of three different forms of the verb is from Palauan (Malayo-Polunesian, Palau). I have indicated morpheme boundaries for convenience:

(53)	Present Middle	Future Participle (conservative)	Future Participle (innovative)	gloss
	mə-ˈdaŋəb	dəˈŋob-l	dəŋəˈb-all	‘cover’
	mə-ˈteʔəb	təˈʔib-l	təʔəˈb-all	‘pull’

The set of surface forms for the word for ‘cover’ is [daŋəb, dəŋob, dəŋəb]. The accent ˈ denotes stress on the following syllable. We can assume that stress is not part of the underlying presentation in Palauan (Chapter 7).

Which of these is underlying? If we compare the verb with the verb for ‘pull’, it cannot be the Future Participle conservative or the Future Participle innovative, since it is not clear how the first schwa would turn into an [a] in one verb and into an [e] in another verb in the Present Middle. But it cannot be the Present Middle either, as in that case, it is not clear why the Future Participle conservative has an [o] in the verb ‘to cover’ and an [i] in the verb ‘to pull’.

Thus neither of these three representations is a plausible candidate for the underlying form. We have to observe that basically the first vowel is unpredictable in the Present Middle whereas the second vowel is in the Future Participle conservative. Because underlying representations contain all information that is not the product of regular phonological processes, the lexicon is the storage of unpredictable linguistic information, and from this point of view it makes sense to say that the underlying forms for ‘cover’ and ‘pull’ are /daŋob/ and /teʔib/ respectively. The phonological process then can look as follows:

(54) V (any vowel) → ə when not stressed

(54) is a typologically very common process in the phonologies of the world’s languages, called *reduction*, as we will see in Chapter 7. Since at least one of the two vowels of the stem will end up being unstressed, it will give as a result that the underlying forms never surface as such. **reduction**

This means that the underlying forms in Palauan are to some extent ‘abstract’ — we cannot observe them directly, but only deduce them from the patterns in which the morphemes occur. There has been a lot of debate within phonology as to how ‘abstract’ we should allow our representations to be. The Palauan representations are still moderately abstract: every aspect of the underlying representation shows up somewhere directly, they just do not occur together in one surface form.

It is more difficult to find cases where we do need to posit e.g. a consonant or vowel that never really shows up anywhere. Cases like this have been proposed, for instance for the analysis of so-called *h aspiré* in French. Certain adjectives have an ('underlying'!) consonant that shows up only before a vowel-initial word and not before a consonant-initial word:

(55) a. *peti*[t] *ami, premie*[r] *ami, bo*[n] *ami* 'small/first/good friend'
b. *peti*[∅] *camerade, premie*[∅] *camerade, bo*[∅] *camerade* 'small/first/good friend'

The process is similar to the one we have seen for Tibetan above: in a sequence of two consonants, one of them (in this case the first one) is not pronounced. In Chapter 5, we will see some complications to this simple pattern, but for now it will serve our purposes, because there are a few words that start with a vowel, but still make the preceding consonant disappear:

(56) *peti*[t] *héros, premie*[r] *héros, bo*[n] *héros* 'small/first/good friend'

The orthographic *h* in these examples is not pronounced; as a matter of fact, French does not have a [h] sound at all, and has not had it for a long time. This is an example of French orthography being conservative. However, some have suggested that facts such as those in (56) indicate that the consonant is still there underlyingly. It would first cause the deletion of the consonant in the adjective before it is itself deleted because [h] is not a sound at the surface in French.

Not everybody agrees with this type of analysis. Alternatives that have been suggested are that words like *héros* are marked as exceptions to the derivational process deleting a consonant, or even that whole adjective-noun combinations are stored in the lexicon with their pronunciation, so that there is no 'deletion' going on at all.

3.3 Interaction of processes

The action within a given language is of course not restricted to just one operation. Several thing might happen at the same time.

We could call the set of all processes of a language its 'grammar'. Chapter 6 will be devoted in detail to the possible computational structuring of this grammar. For now, we observe that if we have different processes, we sometimes should pay attention to how they interact.

Of course, proceses do not always interact. We have discussed above that Kimbundu has vowel harmony, i.e. a process of adapting features from one vowel to the next. It also has a process of nasal assimilation of the consonants in the same suffixes:

(57)

ŋg-a-bit-ile 'I passed'	ŋg-a-bet-ele 'I beat'
ŋg-a-batu-is-ile 'I caused to cut'	ŋg-a-bet-es-ele 'I caused to beat'
ŋg-a-kin-ine 'I danced'	ŋg-a-tom-ene 'I stuck'
ŋg-a-tum-ine 'I sent'	ŋg-a-net-ene 'I had been fat'

The former two examples are repeated from example (47) above; the indicate the vowel harmony. The lower two exxamples show that if the stem contains a nasal, the consonant in the suffix also becomes nasal — this is sometimes called nasal harmony. It also shows that this happens regardless of the quality of the stem vowel. Vowel harmony and nasal harmony are thus independent from one another.

There are some cases where an ordering seems to be even more relevant. A famous case is found in Tiberian Hebrew (Semitic). The first process this language has is deletion of certain unstressed vowels. For instance instead of *saliqu* 'they went up', we find [sliqu]. The second process is one of *spirantisation*: plosive become fricatives after a vowel. For instance, the word for 'he wrote' is underlyingly something like /kaːtab/, but is [kaːθav] on the surface: both /t/ and /b/ occur after a vowel, so they become corresponding fricatives on the surface. **spirantisation**

Now consider the following form:

(58) kaːθvu 'they wrote'

In this case, [v] is the spirantised reflex of /b/ in spite of the fact that it does not occur after a vowel. The reason is that this vowel was there underlyinglys, but it has been deleted by the first process. We thus can only understand this form by assuming that *first* we applied spirantisation to it, and *then* deleted the vowel that triggered the spirantisation. The conditions of spirantisation are thus not 'surface-true', we have to look more abstractly. From a historical point of view, we can of course understand these references to temporal ordering; it is a hotly contested issue within phonological theory whether synchronic phonology has a similar ordering of processes.

3.4 Alternatives models

The preceding pages of this chapter have presented a fairly classical picture of the organisation of the lexicon and the phonological grammar and the way in which they operate together. Like any model in serious disciplines, this model is subject to constant criticism and refinement. In this section, I briefly review some of the main issues.

What holds most of these criticisms together are objections to the algebraic nature of the standard theory as we have presented it in this Chapter. It is a theory that is based on abstract discrete rules operating on discrete elements. The lexicon in this alternative view is a storage of forms that are more detailed in the phonetic detail and therefore more 'concrete' than the standard view. Also, it is sometimes suggested that because many forms are stored, including many forms of the same word and including derived and inflected words, it is not necessary to set up active phonological rules. Instead, the part of human memory that we call the lexicon could be organized in such a way that related words are stored somehow 'close' to each other, where relatedness would not just be defined in terms of morphological relatedness, but also on phonological similarity (how much sounds do the words have in common) and possibly on frequency.

These alternatives thus are less 'algebraic', they propose representations and operations that are less absolute (and therefore more gradient). Such theories require a textbook of their own. We can only scratch the surface here.

Detailed phonological representation

Two important issues have been brought up with the assumption that representations are absolute. The first concerns the fact that phonetic implementation has access to frequency, and the second that phonological rules also seem to have a gradient effect.

The first issue has become known as the *thyme-time difference*. It is known that highly frequent words tend to be shortened considerably, and shortening of an English word starting with voiceless plosive may mean that this plosive has little or no aspiration. The words *thyme* and *time* are usually considered to be homophones, and the standard theory has no other option but to represent them as such; both could be represented by whatever phonological representation corresponds to /tajm/. Also their surface representations should be the same; mainly, aspiration can be added to the voiceless obstruent for phonetic enhancement.

Obviously, these words themselves greatly differ in frequency, *time* being of a very high frequency and *thyme* of a very low frequency in the speech of virtually all English speakers; but for the standard model it should not matter. To the extent that frequency is stored in this model in the lexicon at all, it would be assumed to be stored with the phonological representation, which is shared by the two words, and thus would have the frequency of those two words combined.

It turns out, however, that this does not correspond to facts about how people pronounce these words: *time* is systematically shorter and has systematically less aspiration than *thyme* when tested in a corpus of so-called 'natural' speech (in this case a database of American phone conversations). We could now say that this is a matter of phonetic implementation – but that means that the representation the phonetic module receives for the two words is a different one, and that goes against the idea that phonetics only sees the phonological representation.

Exemplar Theory

Results such as these, showing the importance of frequency and of phonetic detail, have led many linguists to embrace an alternative model of phonology called *Exemplar Theory*. In this model, speakers do not store an abstract representation of a word, consisting of features, etc.; rather they store the precise sound shape of a word whenever they hear it, together with some other information, such as the meaning but also the context in which the word was heard; and they do so in a 'cloud' of such exemplars, where two words that are closer to each other in form, meaning, etc., are also stored closer to each other.

Phonology in such a model is just a form of memory retrieval: whenever we speak, we choose the form that seems most appropriate for the context at hand (this may mean that we take the average of a few forms that we have heard used in more than one context). There is no real phonological calculation going on. Alternations like the ones we have seen in this chapter are all due to historic change.

Frequency effects follow from the model in the sense that more frequent words have bigger, more densely populated clouds. Also, frequency might affect the range of change. As to the phonetic subtlety, or gradience, of such changes, they can also be accomodated for. (Obviously, humans do not have infinite memory, so it is not possible to store an infinity of detail, but they do have a lot of memory, so it is definitely possible to remember more about an individual words than their phonological features.)

More arguments have been adduced in favour of such a model. It turns out, for instance, that the pronunciation of the vowel in (short, monosyllabic) words depends on 'neighbourhood density'. The *neighbourhood* of a word are the words which are phonologically similar to it (*man* has in its neighbourhood the words *ran, mac, mine*, since these words differ only in one segment from it) If this neighbourhood is dense — if it contains many words — people tend to make a more 'extreme' version of the vowel. Again, this shows that the phonetic implementation of a word can 'see' things which it is not supposed to see. **neighbourhood**

At the same time, there also are quite a few arguments in favour for the traditional model. From a number of experiments it appears that there is a 'phonological buffer': phonological structures are first built up before they are interpreted phonologically. This explains speech errors, but also the fact that people seem to take time to compute predictable information, rather than retrieving everything from lexical storage, and that "assembly of longer phonological plans takes longer than assembly of shorter plans".

Scholars are therefore trying to integrate the two models in some way. It has been claimed for instance, that we can distinguish logically between two types of knowledge: *declarative memories*, i.e. lexical/phonological representations ("I know *that* the word *man* consists of one stressed syllable") and *procedural memories*, i.e. phonetic know-how ("I know *how* to pronounce /m/: by closing my lips and then letting the air escape through the nose"). It could be that both kinds of memories play a role in language and speech: the standard models of phonology then describe declarative memory, whereas Exemplar Models describe the procedural memory.

It would mean that there is some overlap between the two. This is in itself not desirable in any theory (as it runs against the well-known theoretical principle of *Occam's Razor*), but we do know that biological systems are often messy and redundant.

Incomplete neutralisation

One of the aspects of phonological representation that we probably want to keep, in spite of all the phonetic detail that seems sometimes warranted, is the notion of *binary contrast* (which we encountered already in section 2.2): we have seen already quite a lot of evidence that within a language with voicing, sounds will be either voiced or voiceless and there are no 'in between' sounds.

Such binary distinctions might sometimes be lifted, for instance in final devoicing, as we have seen above. This is called *positional neutralisation*: the contrast between two sounds is 'neutralised' in certain positions. In German, there is thus allegedly no difference between voiced and voiceless sounds at the end of words, between *Rad* 'wheel' and *Rat* 'rat'. **positional neutralisation**

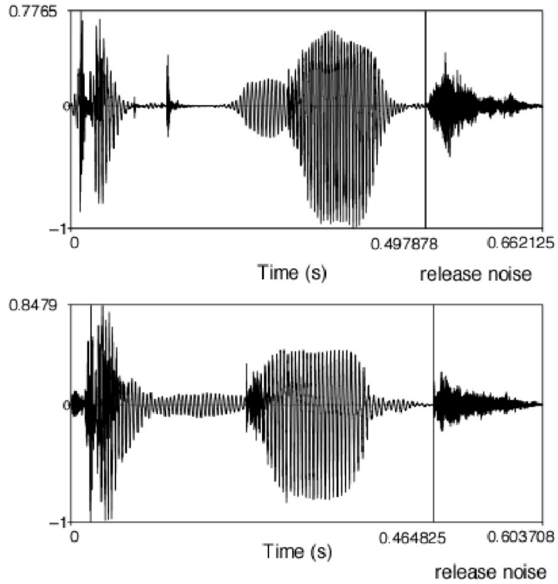
However, careful experimental investigation in a number of languages with final devoicing has shown that under laboratory conditions, speaker do make a difference between words ending in a devoiced and words ending in an underlyingly voiced segment: the former still contains a trace of the original voicing. Furthermore, listeners seem to be able to pick up on this: even though German speakers would typically deny hearing a difference between

Rad and *Rat*, they typically score (slightly, but significantly) above 50% when invited to guess.

In some cases, such as in German, the incomplete neutralisation may be attributed to spelling differences: underlyingly voiced plosives are spelled with letters reflecting their voicing (like the d in *Rad*), and speakers in the experiment might respond to that. However, this cannot be the whole story, as the effect has also been found in Catalan, in which the spelling does reflect devoicing.

Another factor that might play a role in this case are paradigmatic effects: the fact that the final obstruent in *Rad* does correspond to a voiced [d] in the plural might have an effect. In an Exemplar Theoretic model, these words would be somehow linked in the lexicon, and when one pronounces one form, the other form is also activated, influencing this interpretation.

Interestingly, the effect has been found even for nonce verbs. For instance in Dutch, researchers found the following differences between *ik duut* and *ik duud*; *ik* is the first person singular pronoun in Dutch, but neither *duut* nor *duud* exists, so that affixed forms cannot really be activated, except maybe by analogy.



(59)

In the case of Dutch, however, forms ending in voiced obstruents in the first person singular are often past tense morphemes and the slight voicing might be a way to keep this information. Such effects may also be argued to follow from some extension of Exemplar Theoretic models, in which morphemes do not really play a role: words are stored as a whole, and linked to each other with sound and meaning correspondences to other words. There are thus no independent 'morphemes' which are combined by morphology, just like there are no phonological processes. This means that incidental sound-meaning correspondences (voiced plosives often correspond to past tense), that would never be considered 'morphemes' in a standard model, can be of influence here.

At the same time, this means that sound-meaning correspondences like 'past tense corresponds to voiced plosives' somehow should be encoded in the model. Critics of such models might point out that we need a really large number of different possible correspondences, requiring a lot of computational power. Be all of this as it may, it is clear that there are still several aspects of phonological reality that are hard to capture in our models — which however at the same time also seem to be largely on the right track.

Learning the lexicon

If the analyses presented in the body of this chapter are on the right track this has implications for language learning. One of the many tasks a language learning infant has to face is to distinguish individual words in their language. These words will then have to be stored in their lexicon, in their underlying form. This means that the type of analysis we have performed in this chapter, complicated as it may seem, somehow also has to be performed by the child.

As in the analyses presented here, this means that the learner has to find out that there are alternations going on: that certain forms which are different still refer to the same morpheme. But she also has to find the underlying forms; as we have seen, these two factors interact, as one is dependent on the other.

Productivity is the most important argument that something like this still exist: English speakers know how to form the plural of *wug*, although they have never encountered that form. One might say that this is 'by analogy' to other forms, but that then means one has to learn how analogy works — which is a form of computation.

The learning is further complicated because the underlying form sometimes has to be reconstructed from various surface forms, and processes may be obscured because they are made opaque in their ordering. Learning both of these together thus is no trivial task. And of course at the same time, children also have to learn what are the relevant phonological features and how can these features be combined in this language to segments. Those things can interact with the topic of this chapter as well (e.g. a child learning German should learn that the languages distinguishes between /d/ and /t/, so that a feature [Voice] is active in the language *and* that this feature is delinked in word-final position.)

Quite a lot is known about phonological language acquisition; work here runs along two different lines. The first is empirical work on the way in which children actually learn the words and the phonologies of their language: in what order do they learn words and lexical representations. At the same time, we also dispose of computational models, showing how computer programmes can learn a phonological system involving both the underlying forms and the phonological grammar given a realistic input.

It is clear from this work that a lot of rote learning is happening at all stages: learners store complete forms, also if they are morphological complex. But it is also clear that at some point there is some abstraction going on.

3.5 Exercises

1. Could loanword phonology shed more light on the case of complementary distribution in Southern Kikongo? Discuss.
2. In a language with final devoicing, one underlying representation may lead to two different surface representations, and one surface representations may correspond to two underlying representations. Give examples of these. (You may make these up if you do not know any language that has final devoicing well enough.)
3. There is a similarity between the Farsi data described in (45) on page 51 and those of French in section 4.3. Describe this similarity.
4. What are the most likely underlying forms for the prefixes in the following Russian examples? Describe the process that explains the alternations.

(Ła) i. ot-jexat^j 'to ride off'
ii. ot-stupit^j 'to step back'
iii. ob-brosit^j 'to throw aside'
(Łb) i. s-jexat^j 'to ride down'
ii. s-prosit^j 'to ask'
iii. z-delat^j 'to do'
(Łc) i. pod-nesti 'to bring (to)'
ii. pot-pisat^j 'to write'
iii. pod-ʒetʃ 'to set fire to'
(Łd) i. iz-lagat^j 'to state; set forth'
ii. is-kl^{j}uʃat^j 'to exclude; dismiss'
iii. iz-gnat^j 'to drive forth'

5. Two pairs of sounds are in complementary distribution in the following Tohono O'odham (Uto-Aztecan, United States / Mexico) dataset. Which four consonants are these? What is their distribution

- *čuagia* 'net bag', *tatai* 'tendon', *čučul* 'chicken', *ǰuni* 'dried cactus', *čigitog* 'to think', *činig* 'to move', *tatal*tendon, *činig* 'to move', *tamš* 'gums', *tohnto* 'degenerate', *ǰigos* 'storm', *ǰusukal* 'lizard', *dakpon* 'to slip', *ǰiwhiadag* 'arrival', *čukma* 'dark', *čilwin* 'to rub', *tokih* 'cotton', *doʔag* 'mountain', *daswua* 'to pile', *todsid* 'to frighten', *čiposid* 'to brand', *ǰuhki* 'rain'

6. Look at the following data in Spanish. Which are the underlying forms? Explain.

- *dama* 'lady'
- *daðo* 'given'
- *deðo* 'finger'
- *usteð* 'you'
- *donde* 'where'
- *de ðonde* 'from where'
- *la ðama* 'the lady'

7. Consider the following data from Mallorcan Catalan (Indo-European, Spain). Which are the underlying forms and what is the process that might be going on here?

   (Ła) puk 'I can'
   (Łb) pɔt 'he/she can'
   (Łc) sap 'he/she knows'
   (Łd) pɔk 'little'
   (Łe) set 'seven'
   (Łf) kap 'no'
   (Łg) pubːeɲa 'I can wet'
   (Łh) pɔpːan 'little bread'
   (Łi) pudːɔrmir 'I can sleep'
   (Łj) satːot 'he/she knows everything'
   (Łk) sakːontr 'he/she knows how to count'
   (Łl) pokːomprar 'he/she can buy'

8. Turkish has a process of (optional) vowel epenthesis. Below are a few examples. Describe why these vowels are chosen as the epenthetic ones.

   (Ła) /kojn/ → [kojun] 'bosom'
   (Łb) /hykm/ → [hykym] 'judgement'
   (Łc) /metn/ → [metin] 'text'
   (Łd) /sabr/ → [sabɨr] 'text'

9. Russian has final devoicing, witness examples such as the following:

   (Ła) sleda - slet 'track (GEN./NOM. SG.)'
   (Łb) kniga - knik 'book (GEN./NOM. PL.)'
   (Łc) guba - gup 'lip (GEN./NOM. PL.)'

   Describe the interaction between this process and the one you have found in exercise 4, based on the following examples:

   (Ła) pojezda - pojest 'train (GEN./NOM. SG.)'
   (Łb) vizga - visk 'squeal (GEN./NOM. PL.)'
   (Łc) izba - isp 'hut (GEN./NOM. PL.)'

10. For each of the following forms in English, discuss whether or not it is plausible that they are related by phonological computation:

    (Ła) *three-third*
    (Łb) *electric-electricity*
    (Łc) *hated - called* (past tense suffix)
    (Łd) *most apples - mos[∅] people*
    (Łe) *thunder-detonation*

11. Give the underlying form of the Tagalog roots in the following table:

bukas	*buksin*	*buksan*	'open'
kapit	*kaptin*	*kaptan*	'embrace'
tubos	*tubsin*	*tubsan*	'redeem'
opos	*upsin*	*upsan*	'open'
posod	*pusdin*	*pusdan*	'open'
bata	*bathin*	*bathan*	'suffer'
bili	*bilhin*	*bilhan*	'buy'
dipa	*diphin*	*diphan*	'open'
polo	*pulhin*	*pulhan*	'ask for trifles'
pujo	*pujhin*		'saddle bag'
banig	*baŋgin*	*baŋgan*	'mat'
damit	*damtin*	*damtan*	'clothe'
ganap	*gampin*	*gampan*	'fulfill'
putol	*putlin*	*putlan*	'cut'

Sources and further reading

Section 3.2. Finding patterns of phonological structure is the focal point of many older and newer textbooks in phonology. See for instance Kenstowicz and Kisseberth (1979); Halle and Clements (1983); Gussenhoven and Jacobs (2005); Hayes (2009); Peng (2013); Kennedy (2018). A nice introductory article on the importance of underlying representations for our understanding of phonology is Hyman (2018); also Cole and Hualde (2011) is a valuable source, and has been helpful in constructing this chapter. The idea of using complementary distribution is originally — as far as I have been able to find this out — from Pike (1947). Final Devoicing has been studied extensively; in German, but also in other languages. A classic, purely phonological study is Brockhaus (1995). The idea that its neutralisation is incomplete is found among others in Port and Leary (2005); a recent assessment of the research on German is Roettger et al. (2014). The Palauan data are from Schane (1974).

Section 4.5. On enhancement see Keyser and Stevens (2006) and references cited there. The literature on vowel epenthesis is summarised in Hall (2011), and that on consonant epenthesis in Staroverov (2014). Deletion is discussed in Harris (2011). An overview of the discussion on metathesis is in Buckley (2011); the Leti data and their original analysis are from Hume (1998).

Section 3.3. A nice summary of the literature on 'levels of derivation' can be found in McCarthy (2007).

Section 3.4. The original claim that *time* and *thyme* are not homophones is from Gahl (2008); and interesting response is Lohmann (2018), criticizing Gahl's methodology while also showing that her conclusions are correct. The Dutch data on incomplete neutralisation can be found in Ernestus and Baayen (pear).

Autosegmental theory

4.1 Tone

The smallest elements of phonological structure are features, as we have seen in Chapter 2. The next question is how these features are grouped together into larger units, such as segments and, ultimately, words, phrases and spoken utterances. As an approximation of this, we have suggested that segments are bundles (or sets) of features. Words are then strings of such segments. This is based on the assumption that we can still divide the sound stream into discrete time events: one segment comes after the other, although each segment has internal structure.

When we take a closer look at the data from languages around the world, these suggest that the picture is slightly more complex than that. The features behave sometimes as independent from their segments: in those cases, it looks as if one feature is attached to more than one segment, or even that features do not belong to any individual segment at all. Segments thus are not the unique organizing units.

A very influential view of what the organisation of features into larger structures looks like is *autosegmental phonology*. According to this theory, we can see the organisation of speech sounds in language as a musical score: every feature, standing for a part of the speech organ, plays its own part, which is largely independent of all other parts. The only relation between the parts is that they are all attached to one central line, the *skeleton*, which keeps track of the time (left to right corresponds to earlier to later). The elements of the skeleton — which resemble the notion of a segment in certain ways — are usually depicted as x's (we will return to this in section 4.3).

autosegmental phonology

skeleton

An important indication for the correctness of this view is *assimilation*, a process by which two sounds become more similar to each other when they are put in adjacent positions.

assimilation

For instance, the Dutch plural past tense suffix has two allomorphs: -/də/ and -/tə/. The former is chosen after a stem ending in a voiced segment, the latter if the stem ends in a voiceless obstruent (notice that this resembles in an abstract way the pattern for English):

(60) a. *lee[*vd*]en* 'lived'
b. *ma[*ft*]en* 'sleeped'

Why would it be preferable to have a voiced segment next to another voiced segment? If all segments have their own independent choice of voicing, there is no particular reason to expect this preference to hold (rather than, say, a preference for voiced obstruents to be chosen after coronal segments). The answer of autosegmental phonology is: because in this way the two segments can *share* a feature, and this leads to representational economy: we can deal with one feature specification rather than two.

autosegmental behaviour of features

The *autosegmental behaviour of features* — the fact that features behave as if they are independent elements, or segments, in their own right — was first discovered in the 1970s on the basis of tone. In many languages of the world, syllables can differ from each other just by being pronounced at a different pitch: tone thus can be a feature in such languages. Many African and East Asian languages use tone, but so do languages in other continents; it has been estimated that over 60% of all languages in the world use tone to build minimal pairs.

Chinese is a well-known example, in which we distinguish words by their tone only. These tones are marked by accents on the vowel letter in the transcriptions. Mandarin Chinese has four different tones (and some Chinese dialects are claimed to have even more): a mid tone, a high tone, a rising tone and a low tone:

(61) a. *dā* 'to hang over; to hover' (mid tone)
b. *dá* 'to answer or reply' (high tone)
c. *dǎ* 'to beat; to fight' (rising tone)
d. *dà* 'big; huge; great' (low tone)

Tones are pronounced on vowels, and for this reason I have treated them as vocalic features in Chapter 2. Yet tones are demonstrably also independent of their vocalic hosts. The following data from Kikuyu (Bantu, Kenya) show this. The distribution of tones in the word looks rather messy at first, but they actually come in a very simple and clean pattern if we look at them more closely.

In order to analyse the phenomenon at all, we first have to briefly consider the morphological structure of the Kikuyu verb, which can be described by the following template, disregarding the tones:

(62)

SUBJECT	(OBJECT)	ROOT	TENSE
to 'we'	mo 'him'	rɔr 'look at'	
ma 'they'	ma 'them'	tom 'send'	irɛ PAST

A verbal root such as [rɔr] is preceded by prefixes which express properties of the subject of the sentence, as well as the object when the verb is transitive. It can be followed by several suffixes; one of these is Tense (in this case, expressing that an event took place in the past).

If we combine these morphemes and we study the resulting patterns, it looks at first as if (almost) any morpheme can occur both with a low tone (marked à) and with a high tone (marked á):

(63)

	Subject 'to'	Subject 'ma'
rɔr	tò rɔ̀r ìrɛ́	má rɔ́r ìrɛ́
	tò mò rɔ̀r ìrɛ́	má mó rɔ̀r ìrɛ́
	tò mà rɔ́r ìrɛ́	má má rɔ́r ìrɛ́
tom	tò tòm írɛ́	má tóm írɛ́
	tò mò tòm írɛ́	má mó tòm írɛ́
	tò mà tóm írɛ́	má má tóm írɛ́

You can see that the stem vowel, for instance, sometimes has a high tone, and sometimes a low tone. A native speaker of Kikuyu knows exactly which tone to use in which circumstances; but where does this knowledge come from? How can you know this? If a tone is stored in memory on a specific vowel — as is clearly the case in Chinese — which vowels are that for each of the tones in Kikuyu?

On closer inspection, it is not precisely true that any vowel shows any tone: the vowel of the subject marker 'to' always comes with a Low tone, while the vowel of the subject marker 'ma' always comes with a High tone in (63). Furthermore, the morpheme immediately following the subject marker always has exactly the same tone as the subject marker itself. In some sense, the subject thus seems to determine the tone of the following morpheme. It is as if the tone of the subject marker also gets expressed on the object marker or the stem.

Similarly, we may observe that the final tone of the tense suffix is always high, but that the first vowel has a varying tone: if the stem is *rɔr*, we find a low tone, if it is *tom*, it is a high tone. Thus it seems to be the stem which determines the first tone of the suffix.

We can best understand the Kikuyu tone system if we generalise these observations: the tone of every morpheme shows up on the following morpheme. Every morpheme in Kikuyu thus consists of two separate parts: segmental material on the one hand, and completely independent of that a tone which is realised on the *following* vowel.

We can thus determine that every morpheme has a tone like this:

(64)

to	ma	mo	ma	rɔr	tom	irɛ
L	H	L	H	L	H	H

On the surface, every tone needs to be linked to some vowel, and none may be left 'floating around' (the term *floating feature* is a technical term in autosegmental phonology, describing features that are not linked to a segment), due to the so-called Association Convention of autosegmental phonology: **floating feature**

(65) *Association Convention*: No 'floating' tones are allowed when we pronounce a word, every tone needs to be linked to a vowel. **Association Convention**

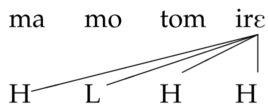
The Association Convention for tones is part of a more general set of requirements on phonological structure, requiring every element in a phonological representation to be linked to the other parts of the phonological structure. In many languages, the tones would be linked to the vowel in their own morpheme, which would obviously also be the most logical option, but in Kikuyu there apparently is a different requirement which is more important than keeping every tone realized in its own morpheme: *alignment* of the tone **alignment** to the end of the word. We can express this by the following constraint on Kikuyu words:

(66) ALIGN-Tone: All tones want to be as close to the right edge of the word as possible, given other conditions of the language.

spreading

In many tone languages of the world, we see the effect of ALIGN-Tone: tones tend to move to the right, which we call *spreading*. There might be a phonetic reason for this: the realization of tones tends to be a little bit delayed in pronunciation, and spreading (rightward) is a phonological reflex of this. In any case, it is a tendency we see in many tone languages.

When ALIGN-Tone would decide things on its own, regulating the realisation of Kikuyu words as it sees fit, it would choose to have the following representation as the best one for *ma mo tom irɛ*:

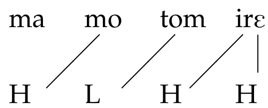
(67)


All tones are linked to the final vowel, and thus maximally aligned to the end of the word. The advantage of this comes at the enormous cost, however, of creating a very complex tonal configuration on this final vowel, and apparently, this is not a price which Kikuyu is willing to pay. In particular, the relation between tones and vowels in this language is very transparently one-to-one. In other words, the Association Convention above can be refined to the following:

Wellformedness Condition

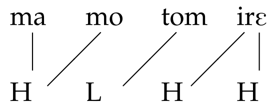
(68) *Wellformedness Condition* (WFC): Every tone in the output representation should be linked to exactly one vowel, and vice versa.

Given strong force of the WFC in Kikuyu — it is not absolute in all languages, as we will see later — the best we can do to maximally satisfy ALIGN-Tone is the following:

(69)


Every tone is now linked as much to the right as possible, without creating illicit 'contour' tones. Notice, however, that there is still one problem: the very first vowel (the one of the subject marker) does not bear a tone at all. There is no way we can solve this problem, paying due respect to all the requirements imposed on the Kikuyu word, and some related Bantu languages would leave it like this in similar situations, creating a toneless syllable.

However, notice that the WFC expresses several requirements at the same time, e.g. 'no tone should be linked to more than one vowel', and 'no vowel should be toneless'. Apparently, the former counts as a stronger violation in Kikuyu than the latter and therefore the following repair is made:

(70) 

Notice that the correct output representation is found in Kikuyu by striking the right balance between several, sometimes conflicting, forces. Other languages might be subjected to similar forces, but their respective weights might cause the balance to be different in some cases. We will go deeper into this computational mechanism and how we can derive the typology of languages from it in Chapter 6.

Contour tones

As we have seen, Kikuyu is very strict in its requirement that vowels can be linked to at most one tone. The idea that tones exist in their own dimension of representation — called a *tier* in autosegmental theory — and that their alignment to tones can be regulated by a set of requirements — called *constraints* — is central to (modern versions of) autosegmental theory.

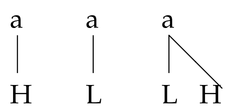
tier

constraints

Another application of the idea of this theory which has proved very useful, is the analysis of so-called *contour tones*. For instance, Margi (Chadic, Nigeria) does not have two but three different tones on vowels: a low tone, a high tone, and a rising tone, and is therefore more comparable to Chinese.

contour tones

Obviously, this situation cannot be described by just two features, [High] and [Low]. In principle, there are two ways of dealing with a situation such as this. We can either introduce a three way featural distinction (e.g. a feature Tone which has as values High, Low and Rising); or we can describe the rising tone as a combination of Low followed by High. Autosegmental analysis advises us to take the latter route, so that we can minimize the number of primitives in our theory (there are only high and low tones, and autosegmental association):

(71) 

For Margi, the advice that autosegmental phonology gives us turns out to point us in the right direction. In the first place, this representation helps us to understand what is going on with tones when morphemes combine into complex words. Consider the following facts concerning the definite suffix *-ári*. The left-hand column represents the underlying shape of the stems to which this suffix is added (ˇa represents the rising tone):

(72)	a.	sál	sál-árì	‘man’
		kùm	kùm-árì	‘meat’
	b.	ʔímí	ʔímy-árí	‘water’
		kú	kw-árì	‘goat’
	c.	tì	ty-ǎrí	‘morning’
		hù	hw-ǎrì	‘grave’

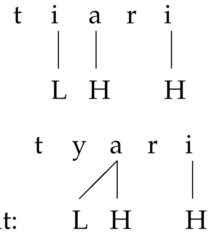
(72a) shows that nothing happens if the suffix is attached to a consonant-final stem. Unlike in Kikuyu, every morpheme keeps its own home base in Margi; apparently the tone of the suffix is high.

(72b) shows that if the stem ends in a high vowel with a high tone, this turns into a glide. Since glides, like all consonants, cannot carry their own tone, it looks as if the high tone disappears.

(72c) shows that something does happen if the stem ends in a high vowel with a low tone. Again, the vowel turns into a glide, but now the tone of the suffix changes to a rising tone. Under autosegmental assumptions, it is very easy to understand this process: the rising tone is a combination of the original low tone of the stem and the high tone of the suffix:

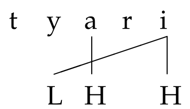
(73) a. Input:

b. Output:



The reason why this happens can be seen as an interaction of the impossibility of the glide to carry the tone, and the wish of the tone to be linked to some vowel. Notice, by the way that this is always the vowel which is closest to the tone in some intuitive sense. In particular, we will not find the following structure (the representation for *tyárǐ*):

(74)



The reason why we do not find this, is that there is yet another very hard constraint on autosegmental representations:

NoLineCrossing (75) *NoLineCrossing*: Association lines may not cross

Different from all other constraints we have seen so far, NOLINECROSSING is hard-wired into every known grammar: languages cannot fiddle with it and create exceptions like with the other constraints we have seen so far. The reason for this presumably has to do with the interpretation of autosegmental representations. We are dealing in this case with two lines (traditionally called *tiers* in the theory): one line on which we have the tones, and another line on which we have the relevant vowels.

Each of those tiers represents a timeline: if element A stands before element B on a tier, this means that the pronunciation of A precedes the pronunciation of B. Thus, in (74), the realisation of the low tone will always precede that of the high tones.

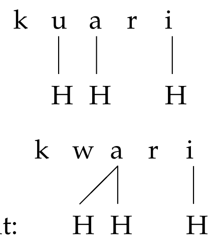
If we think about our representations in this way, it stands to reason that association of an element X to an element Y means that the realisation of X overlaps with that of Y in time. Thus the pronunciation of the low tone in (73a) will happen during the pronunciation of the /i/.

Given all of this, (74) defies ordinary logic: the low tone precedes the first high tone, but it is also realised during the pronunciation of an [i] which follows the [a] with which the low tone is associated. In other words, the pronunciation of the low tone will also follow the pronunciation of the high tone. This is logically impossible: α cannot at the same time precede and follow β (except if they overlap, but that is not the case here).

We can thus conclude that grammars can entertain all kinds of representations, including those which are not completely well-formed (because they display contour tones, or floating tones, or toneless vowels); but they will never entertain possibilities which do not make any sense at all. That is the reason why (75) is never violated in any language.

Another remark to be made with respect to (73), is that this raises the question what is exactly the output representation for e.g. *kwárí*. We may assume that the high tone of the stem is deleted, but it is also logically possible to assume the following:

(76) a. Input:
b. Output:



This would make the high vowel and low vowel stems exactly parallel. Whether or not we accept this, seems to be a matter of taste. Scholars who like the parallelism will readily accept this; others will point out that there is no empirical difference between a segment linked to two tones and one linked to one tone, and that we should therefore go for the simplest representation. The matter is hard to decide.

We quickly look at yet another argument in favour of the representation of rising tone as a sequence LH, to be derived from the underlying structure of stems in Margi. Bisyllabic stems in Margi come in three flavours: some of them have two low tones, some of them have two high tones, some of them have a low tone followed by a high tone. Monosyllabic stems similarly exist in three variants: some have a high tone, some a low tone, and some a rising tone. Under the autosegmental assumption, we can unify these by assuming that there are only three tonal templates in Margi: H, L, and LH:

(77)

	H	L	LH
bisyllabic	*ndábyá* 'touch'	*gə̀rhù* 'fear'	*pə̀zú* 'lay eggs'
	tə́dú 'fall down'	*dzàʔù* 'pound'	*ngùrsú* 'bend'
monosyllabic	*tsá* 'beat'	*dlà* 'fall'	*hǔ* 'grow up'
	sá 'go astray'	*ghà* 'reach'	*və̌l* 'fly'

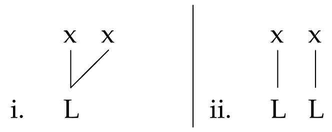
Notice that this means that, even though Margi allows (rising) contour tones, it still only does this as a last resort: only because otherwise a tone would be lost (as in the gliding cases just discussed) or because it is the only way to express a tonal template. A bisyllabic word **gə̀rhǔ* is still not allowed, since it contains an 'unnecessary' rising tone. We thus cannot say that the wellformedness condition does not play a role at all; it just seems to be less stringent in Margi.

Multiple tones vs. multiply linked tones

We thus observed that Margi has three types of disyllabic words:

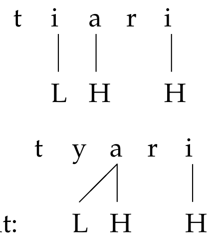
1. The first syllable has a low tone, the second syllable has a high tone.
2. Both syllables have a low tone.
3. Both syllables have a high tone.

The representation of the first of these is straightforward in autosegmental terms, but for the other two, we logically speaking have two options, which I will illustrate on the low tone example:

(78)


There is one reason for assuming that the representation in (i) is the 'real' one: this allows a more uniform description of disyllabic and monosyllabic forms; recall that the latter had three tones: low, high and rising; there is no reason to assume that low toned monosyllabic stems have two (low) tones.

This reason is not very strong, but fortunately there are other arguments and they point in the same direction. First, remember what happened to monosyllabic stems when their vowel would get lost:

(79) a. Input:
b. Output:


And now consider the fate of bisyllabic low toned stems in the same circumstances:

(80)	làgù	làgwári	'road'
	màlà	màlári	'woman'

No rising tones are created in this case. This is much easier to understand if we assume the representation in (78a) for stems of this type than under the assumption of (78b). Under the latter representation, we would not expect any difference with monosyllabic stems: if the final vowel turns into a glide, the second low tone will go and try to find a new host on the suffix vowel, creating a contour tone in the process. But under (78a) there will be only one low tone on the stem, and this low tone does not run the risk of becoming floating, since it can still be linked to the first vowel.

We could now wonder whether there are languages which have both (78a) and (78b) in their inventory of phonological structures. It has been a claim of autosegmental phonology that this is not possible; phonological structures would be subject to the so-called *Obligatory Contour Principle*:

(81) *Obligatory Contour Principle* (*OCP*)
Adjacent identical tones are disallowed.

Obligatory Contour Principle
OCP

The OCP allows tonal tiers like 'H', 'HL', 'L', 'LH', LHL', etc.; but it disallows structures like *'HH' or *'HLHLLH'. If we have two vowels in a row which are pronounced at the same pitch, there is only one option: these vowels are linked to the same tone.

The OCP is again an example which not all languages have to strictly obey (there is no logical necessity to it), but it can be demonstrated that it shows up in many languages, although sometimes under different names, because not all linguists have been aware of the generality of the phenomenon.

Meeussen's Rule in Bantu

In the traditional tonology of Bantu languages, an OCP related rule is called *Meeussen's Rule* (after the Belgian Bantuist Achilles Emile Meeussen, 1912-1978). This rule can be illustrated by the following example, from Kirundi (Bantu, Burundi):

Meeussen's Rule

(82) a. nà-rá-zì-báriìrà (I-PAST-them-to sew) 'I was sewing them'
b. nà-rá-bàriìrà (I-PAST-to sew) 'I was sewing'

In (82a), the high toned tense marker *rá* and the stem *báriìrà* , which also starts with a high tone, are separated by a low tone agreement marker. Nothing happens here; we may assume that this form represents the underlying state of affairs quite faithfully. In (82b), on the other hand, the tense marker and the stem are adjacent. As a result of this, the second high tone has to go.

It is quite obvious that Meeussen's Rule describes an OCP effect: two high tones which are adjacent are not allowed. The way to solve the OCP problem here is to turn one of the two 'bad' tones into a 'good' tone, giving an alternation of high and low tones.

Here is another example of the same phenomenon in a different Bantu language (Shona; Zimbabwe/Zambia):

(83)	mbwá	'dog'	né#mbwà	'with a dog'
	hóvé	'fish'	né#hòvè	'with a fish'
	mbúndúdzí	'army worms'	sé#mbùndùdzì	'like army worms'
	bàdzá	'hoe'	né#bàdzá	'with a hoe'
	bénzíbvùnzá	'inquisitive fool'	né#bènzìbvùnzá	'like an inquisitive fool'
	Fárâì	(name)	nà#Fárâì	'with Farai'

The examples show — among various other things — the following: Meeussen's Rule (i) applies between (certain) clitics and stems, (ii) if the clitic has a high tone and (iii) the stem starts with a high tone. Of interest are the cases in which the stem starts with more than one high-toned syllable. It turns out that in those cases, *all* of those syllables become low toned, even though it would be sufficient for Meeussen's Rule if we would only change the first one (witness what happens to forms such as *né#bàdzá*, where it also not necessary to change the second high tone of the stem).

This behaviour of low toned words can be understood if we assume that two adjacent syllables in the underlying representations pronounced at the same pitch are associated to the same tonal autosegment. If this tone has to change, all vowels attached to it will be pronounced differently:

```
                 n e    h o v e
                   |      |/
(84)  a. Input:    H      H

                 n e    h o v e
                   |      |/
      b. Output:   H      L
```

Interestingly, there are certain sequences of high tones which do not change; but there is always an extra morpheme boundary involved in those cases. For example, we can 'stack' clitics in Shona, leading to sequences such as:

(85) *sé#nè#hóvé* 'like with a fish'

Notice that it is only the tone of *ne* which changes in this case. This high tone is *not* the same as the stem tone. Therefore the latter does not automatically change with the former.

Under the assumption that Meeussen's Rule is an instance of the OCP, the latter principle takes two different effects in Shona:

1. It disallows sequences of the same tone in underlying forms, preferring multiply linked tones instead.
2. It disallows sequences of high tones on the surface, solving apparent problems not by spreading, but by changing one tone from high to low.

There is a third way in which the OCP is operative in this language: it can also block rules from applying. This is true in particular for a rule spreading a high tone from the end of one word to the first syllable of the next word:

(86)

zvìròngó	'water pots'
zvìnà	'four'
zvìròngó zvínà	'four water pots'
Chìpó	(name)
àkàbìkà	'and then he cooked'
Chìpó ákàbìkà	'and then Chipo cooked'
ndàkáténgá	'I bought'
bàdzá	'hoe'
ndàkáténgá bàdzá	'I bought a hoe'

The last example shows that the spreading of the high tone does not occur if the second syllable of the second word already has a high tone. Spreading here would result, again, in a sequence of vowels linked to different high tones, and apparently, this is disallowed.

All in all, the OCP can thus have three effects in Shona:

1. It disallows certain underlying structures (by way of a **Morpheme Structure Constraint**)
2. It can **trigger** certain processes (H→L in clitic structures)
3. It can **disallow** certain processes (H spreading)

This multifunctionality is quite typical for constraints on phonological representations: they can work out in different ways, just bluntly disallowing a structure altogether, or specifying a repair when it arises. Not every constraint will have all of these effects in every language, but very often there is a range of ways in which languages can work towards satisfaction of their constraints.

4.2 Autosegmental representations outside of tone

The OCP gives us a good handle on extending autosegmental ideas to areas beyond tone. In many dialects of Dutch (Indo-European, the Netherlands/Belgium/Surinam), the default allomorph of the diminutive suffix is *-ke* ([kə]). The following example is from Bergen Dutch:

(87) *vrouw* 'woman' - *vrouwke* 'woman-DIM' [vrɑukə]

However, if the stem ends in a velar obstruent, we find the form *-ske* ([skə]) instead (the second example also illustrates umlaut, which is irrelevant for our purposes):

(88) a. *vlieg* 'fly' - *vliegske* 'fly-DIM' [vlixskə]
b. *boek* 'book' - *buukske* 'book-DIM' [bukskə]

This can be understood as follows: bare addition of -[kə] to the stem would result in an OCP violation on the feature [velar]:

```
        v  l  i  x    k   ə
                 |    |
(89)           [vel] [vel]
```

Inserting a segment with a different place of articulation — such as coronal [s] —, solves the problem: the two segments with the 'bad place' are no longer adjacent.

A famous case of a non-tonal OCP effect is the interaction of *Lyman's Law* with the *Rendaku* rule in a certain class of words in Japanese (Japonic, Japan). The latter 'rule' wants the second element of a compound to be a voiced segment even if it is voiceless underlyingly; the former expresses the condition that there is no other voiced segment elsewhere in the word (90):

Lyman's Law

Rendaku

(90)	*tama*	'ball'	*teppoo+dama*	'bullet'
	sono	'garden'	*hara+zono*	'flower garden'
	taba	'bundle'	*satsu+taba*, **satsu-daba*	'wad of bills'
	sode	'sleeves'	*furi+sode*, **furi-zode*	'long-sleeved kimono'

Clearly, Lyman's Law — which despite its name was first discovered by Motoori Norinaga in the 18th century — could be stated as a specific instance of the OCP:

(91) *Lyman's Law* (OCP style): Avoid two voiced obstruents within the same word.

The claim is thus that Lyman's Law blocks Rendaku in Japanese in the way in which the OCP blocks high tone spreading in Shona.

There are various interesting problems connected to this. Most important among these is the issue that apparently vowels and sonorant consonants do not count for the OCP; they are, as it were, invisible. The first examples in (90) demonstrate this clearly. The standard way of understanding this is by assuming that these segments simply do not have a link to any [±voice] feature: they are *underspecified* for that feature. The reason for this is that they do not contrast for this feature: there are no minimal pairs of words where one has a voiced sonorant or vowel, and the other one a voiceless one. Implicit in our analysis of tone above was, by the way, similarly that consonants are underspecified for tones. Again, the reason for this is that the Bantu languages we discussed simply do not distinguish consonants from each other by tone.

underspecified

Vowel Harmony

Another domain to which autosegmental analysis has been applied with considerable success is vowel harmony, a phenomenon that can be found in many languages of the world, albeit in different versions, and of which we have already seen examples in Chapter 3.

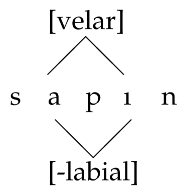
vowel harmony

In a typical *vowel harmony* language, the set of vowels can be split up into two or more disjoint subsets; all the vowels within one word are then taken exclusively from one subset. In Turkish, we can divide the set of vowels along the round-spread dimension as well as along the front-back dimension, which means we have four subsets. The following gives a general idea of what is going on:

(92)		*nom.sg.*	*gen.sg.*	*nom.pl.*	*gen.pl.*
	'rope'	ip	ipin	ipler	iplerin
	'girl'	kız	kızın	kızlar	kızların
	'face'	yüz	yüzün	yüzler	yüzlerin
	'stamp'	pul	pulun	pullar	pulların
	'hand'	eḽ	eḽin	eḽler	eḽlerin
	'stalk'	sap	sapın	saplar	sapların

We can understand this within the principles of autosegmental phonology by assuming that the features [±velar] and [±labial] can (and should) spread in Turkish:

(93)



The idea is that the phonological properties which are expressed by the harmonic features belong to the word as a whole, and get associated to everything within that domain (more about what a 'word' is in phonology in Chapter 8); but they can 'see' only those things which have a harmonic counterpart, i.e. for which the feature makes sense. Since consonants usually do not have a harmonic sister, talking about e.g. [±round] does not make sense, and therefore consonants do not participate in the harmony.

Some consonants in Turkish do have a harmonic sister, however. Exactly those consonants can therefore participate in the harmony. I concentrate on /k/ here, but similar things can be said aboud /g, l/:

(94)

-back /k̭/	+back /k/
k̭ir 'dirt'	*kır* 'meadows'
k̭el 'bald'	*kul* 'slave'
k̭ör 'blind'	*kol* 'arm'
dik̭ 'upright'	*sık* 'often'
dök̭ 'pour'	*ok* 'arrow'
sak̭in 'calm'	*sıkan* 'warning'
fak̭ir 'poor'	*mika* 'mica'
nek̭tar 'nectar'	*boksit* 'bauxite'
bol 'abundant'	*boḽ* 'cocktail'
kar 'snow'	*k̭ar* 'profit'

/k, k̭/ can also initiate harmonic behaviour themselves; to be precise on epenthetic vowels:

(95)

	careful form	*colloquial form*
'fetters'	pranga	pıranga
'prince'	prens	pirens
'test'	prova	purova
'announcer'	spiker	sipiker
'credit'	kredi	kıredi
'cruiser'	kruvazör	kuruvazör

This can also be seen in words like *kulüp* 'club', and even dialects for suffixes, in which the vowel does not harmonize with the final stem vowel but with the final stem consonant:

(96)

	nom. sg.	*acc. sg.*
'explosion'	infiḽak	infiḽak̭i
'perception'	idrak	idrak̭i
'desire'	ševk	ševkı (in some dialects; older speakers)
'confirmation'	tasdik	tasdikı (in some dialects; older speakers))

The behaviour of these features is thus similar to the behaviour of tone in tone languages: they can spread and be linked to more than one segment.

4.3 The skeleton

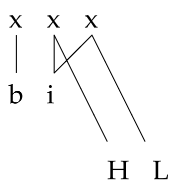
The core of the phonological structure

Until now, we have been quite informal as to the precise structure of autosegmental representations. We know that tones and features such as [Voice] and [Dorsal] can behave like autosegments on independent tiers which are somehow linked to the 'segment', but we have not yet developed a clear notion of how all of these tiers are then organized into a larger structure.

timing tier
skeleton
timing slots

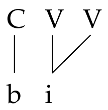
We first discuss the central tier of autosegmental representations, the *timing tier* or *skeleton*. Different from other autosegments, the elements on this tier do not correspond to their own independent (articulatory) instruction. Rather, each of them is represented as a neutral symbol '*x*'; these symbolic units are called *timing slots*, because their most important function is to organize all autosegments into temporal units. They are also sometimes called *x slots*. (Notice that this means that the OCP should not be able to apply to this tier, otherwise we would only be allowed to have 1 segment per word. There is no language which has such a restriction and indeed such a language would be very difficult to use.)

Syllable structure is built on top of these timing slots, and all autosegmental features are linked to them. An autosegmental representation for a (hypothetical) word [piː] with a falling tone would thus be approximately as follows:

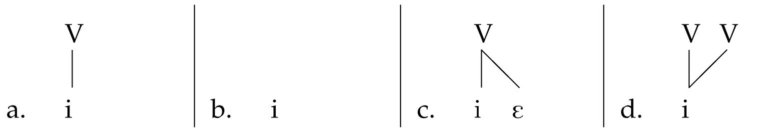
(97)


Several remarks are in order here. In the first place, it may seem as if some of the association lines are crossing in this representation, even though we have argued above that this is absolutely disallowed in the phonology of natural language. The reason for this line crossing is equally trivial, however: we are drawing a three-dimensional structure in two-dimensional space. The tonal tier is in a different dimension from the tiers with segmental information; therefore the lines do not really cross. The logical problem connected with line crossing which we discussed before therefore does not arise — this structure is perfectly legitimate.

Another thing we should note is that (97) is still overly simple. We have conflated many autosegmental tiers by just writing /b/ and /i/. For now, I will also no longer consider the tonal tier. Furthermore, I will stick to a further notational simplification, to sometimes already mark *x*'s as being a consonantal (C) or vocalic (V). All of this means that the word [biː] gets the following representation in our current discussion:

(98) 

From our discussion of autosegmental representations, we now know that we can expect the following variants, for instance for the vowel /i/:

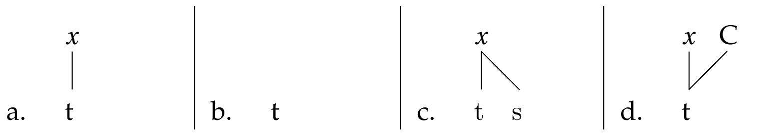
(99) 

Indeed, all of these structures are attested: (99a) gives a regular (short) vowel; (99b) gives a 'floating' vowel (we will see what floatingness means for segments in Section 4.3); (99c) gives a *diphthong*; and (99d) gives a *long vowel*.

See p. 82

diphthong
long vowel

We should also expect that (99a) (with a one-to-one association) is the regular case, which every language has. This is indeed what we find. For consonants we can set up the same set of structures:

(100) 

(100a) gives a regular (short) consonant; (100b) a floating consonant; (100c) a doubly articulated consonant; and (100d) gives a long consonant.

In what follows, we will see examples of the most important (and most surprising) examples of these from a variety of different languages, both for consonants and for vowels.

Long vowels in Finnish and in Germanic

The superlative suffix for nominative singular adjectives in Finnish is *-in*. If we add this suffix to a vowel final stem, the stem vowel gets lost:

(101)

vanha 'old' [vanha]	*vanhin* [vanhin]
köyhä 'poor' [køyhæ]	*köyhin* [køyhin]

The last example shows that it is only the final vowel that gets deleted. At first sight, long vowels are exceptional: they do not get deleted, they get shortened instead:

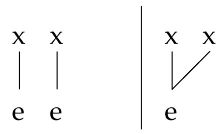
(102)

tervee- 'healthy' [terveː]	*tervein* [tervein]
rakkaa- 'beloved' [rakːaː]	*rakkain* [rakːain]

The behaviour of long vowels is hard to explain under the assumption that they would be carrying for instance a feature [long]. Things become clearer if we consider stems ending in two vowels. In such cases, the first vowel does not get deleted:

(103) *tärkeä* 'important' [tærkeæ] | *tärkeä* [tærkein]

We can unify the 'exceptions' in (102) with the piece of data in (103) if we assume that long vowels consist of two short vowels in a row. The behaviour of *tervee* then becomes completely parallel to the behaviour of *tärkeä*. This however would still leave us with two options for an underlying representation:

(104) 

It is hard to decide on independent grounds, within Finnish, which one of these two representations is the correct one. The language has vowel harmony, but this affects long and short vowels, and diphthongs or vowel sequences all alike.

In some other languages, we may find evidence that the representation on the right is the correct one. One piece evidence comes from a phenomenon called **compensatory lengthening**. An instance from this comes from the history of English. Compare the following Old English words with their Dutch or German cognates:

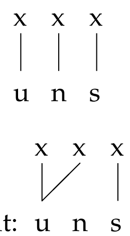
(105)

Old English	*Dutch/German*
gōs	*gans* (Dutch)
ōþer	*ander* (Dutch)
sōfte	*sanft* (German)
fīf	*fünf* (German)
ūs	*ons* (Dutch)

The Old English words all have a long vowel, where the Dutch/German forms have a short vowel followed by a nasal. There is reason to assume that the latter is more faithful to the state of affairs in Proto-Germanic, predating all of the Germanic languages, and that English is the language that has changed.

How can we describe what has happened? Autosegmental phonology gives us a nice tool to provide this description: first, the nasal got lost, i.e. it was delinked from its position on the skeletal (for some reason which we cannot describe at this point yet). After this, the empty position was filled by the preceding vowel. The lengthening is compensatory in the sense that the vowel length compensates for the lost consonant:

(106) a. Input: b. Output:



A few things have been noted again. In the first place, the change we are witnessing is *a priori* nothing like a phonological rule in the sense we have seen them before. We are dealing with a diachronic change, the 'input' in (106) represents some stage in the history of English and the 'output' some other stage; and there is as yet no specific reason to assume that any speaker ever had both of them in his head. Still, also diachronic changes like this may give us some insight in the mental representations of speakers.

Under this assumption, then, we actually see autosegmental phonology at work. If a long *ū* would be nothing but a sequence of two short *u*'s, we would not really understand what was going on: we would have to say that the nasal would have turned into a full copy of the preceding vowel, which would make the representation of this change rather complex.

Compensatory lengthening is found in many of the world's languages. A very well-known case can be found in Turkish. In this language, there is actually a reason to assume that it is a synchronic process and not just the result of language change, because, depending on sociolinguistic and pragmatic factors, speakers can choose to delete or not delete a consonant (to be more precise, one of /v, j, h/. When they do delete, compensatory lengthening follows suit automatically:

(107) a. *kahya* 'steward' [kahja]-[kaːja]
b. *eylül* 'September' [ejlyl]-[eːlyl]
c. *sevmek* 'love' [sevmek]-[seːmek]

It would of course still be possible to assume that the forms with and without consonants are all stored in the lexicon of speakers who show this kind of optionality; but it is not very likely.

Long consonants in Italian

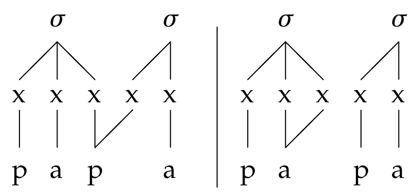
Next to long vowels, we also expect to see long consonants. And indeed, there is at least as much evidence for their autosegmental representation as there is in the case of long vowels.

A famous case comes from Italian dialects in which we find a phenomenon of *Raddoppiamento sintattico* (Syntactic doubling). In the first place, we have to know that most Italian dialects have long consonants (or **geminates** as they are usually called). For instance, there is a contrast between *papa* 'father' [papa] and *pappa* 'porridge' [papːa], which can presumably only be described in these terms. Yet it should be noted that on the surface, the first vowel in the word for father lengthens, whereas the second vowel does not.

The reason for this presumably is that the following is true for Italian (as well as for many other languages):

(108) An x-slot has to follow the stressed vowel.

In order to understand why (108) would need to be the case, we would need to delve deeper into the theories of syllable structure (Chapter 5) and stress (Chapter 7); for now the idea is that stress needs some space within the syllable to be expressed. This space is already available because of the long consonant in *páppa*, but it needs to be filled by the vowel in *pápa* (the accents denote that stress is on the first syllable in both words; σ indicates that segments belong to the same syllable):

(109) 

There are thus two ways of satisfying (108): either by a long consonant or by a long vowel. For some reason, the latter option is not open in the last syllable of the word in Italian: the language does not allow words to end in a long vowel. Therefore, what we find is a doubling of the consonant (which is dependent on certain syntactic factors as well, hence the name):

(110) a. Cittá [sː]anta
Holy city
b. La scimmia aveva appena mangiato metá [bː]anana.
The monkey had just eaten half a banana.
c. La scimmia aveva appena mangiato quáttro [b]anane.
The monkey had just eaten four bananas.

At a sufficiently high level of abstraction, the phenomenon looks a little bit like compensatory lengthening, except that the position to be filled is not caused by deletion of a segment, but by stress. Again, it is hard to understand this without autosegmental representations: why would otherwise the empty position be filled by an exact copy of the following consonant or the preceding vowel?

Floating consonants in French

We now turn from doubly linked structures (long vowels and long consonants) to unlinked structures in the realm of segmental representation: how should we understand 'floating segments', not linked to any timing slot? In particular, we will have a brief look at floating consonants in the phenomenon of *French liason*.

French liason

In this language, final consonants of certain words are subject to a phonologically motivated alternation: they surface before a word starting with a vowel (noted as # V, where # is the symbol phonologists sometimes use for the beginning and the end of the word), but not before a word starting with a consonant (# C) or at the end of a phrase (the double boundary sign ## indicates this):

#

(111)

# V	# C	##
petit ami	*peti~~t~~ camerade*	*il est peti~~t~~*
gros enfant	*gro~~s~~ camion*	*il est gro~~s~~*
un enfant	*u~~n~~ gros enfant*	*il y en a u~~n~~*
premier etage	*premie~~r~~ cas*	*il est le premie~~r~~*

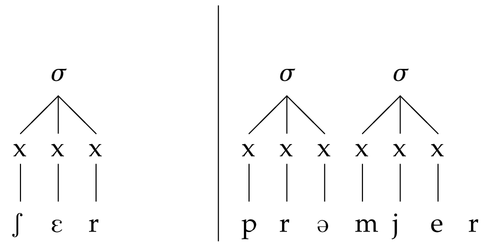
The nature of the consonant that surfaces before a vowel is determined by the preceding word: *petit* always has [t], *premier* always has [r], etc. Thus, these consonants somehow have to be present in the underlying representations of these adjectives.

Further, we have to distinguish the /r/ from *premier* from that of *cher*, since the latter does not alternate, but always surfaces, also before consonants and at the end of a phrase:

(112)

# V	# C	##
cher ami	*cher camerade*	*ça coûte cher*

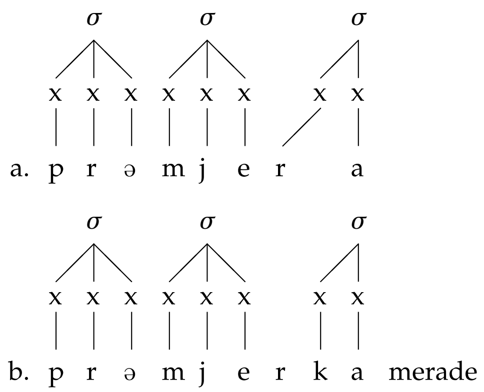
The autosegmental solution is to assume that the /r/ in *cher* is underlyingly linked, whereas the one in *premier* is floating:

(113)


In French, like in many other languages of the world, syllables prefer to start with a consonant rather than with a vowel (in Chapter 5 we will learn that this means that they have to start with an onset): syllables are optimally CV. In parallel to (108) we thus have (114):

(114) An *x*-slot has to precede the vowel in a syllable.

Because words which start with a consonant underlyingly, already satisfy (114), nothing will happen in *premier camerade*: the /r/ will not find an x-slot to be linked to, hence it will not be timed and not pronounced. Yet in *premier ami*, the extra consonant projected because of this requirement, comes to the rescue of the floating consonant.

(115)


Contour segments in Luganda

prenasalized consonants

To round off our discussion of the various autosegmental possibilities of the skeleton, we need of course also to provide evidence for the existence of structures where more than one segment is linked to one timing slot. One piece of such evidence we can find in Luganda, a Bantu language from Uganda. Like many Bantu languages, Luganda has so-called *prenasalized consonants* such as [^{m}p, ^{m}b, ^{n}d, ^{n}t, $^{\text{ŋ}}$g, $^{\text{ŋ}}$k] and a few others. One might think of them as two segments (a nasal and a plosive) but at the same time they behave like one segment, for instance with respect to syllable structure (which we will not discuss).

Furthermore, they are always preceded by a long vowel:

(116) a. *ku siinza* 'to worship'
b. *ku toonda* 'to create'
c. *mu leenzi* 'boy'
d. *ku laba* 'to see'
e. *ku*: *n daba* 'to see me'

The last example shows that the lengthening is not just another instance of a diachronic process, but it corresponds to a productive rule of Luganda phonology. It also shows that if we put a segment /n/ together with a stop, we create a prenasalized consonant.

The autosegmental analysis of this is not too complicated. Apparently a nasal will dock unto the *x*-slot of the following consonant, for whatever reason (maybe because the language does not like to have two consonants linked to independent x-slots in a row, due to some sort of OCP effect). Because of this, the original x-slot of the nasal becomes available, and the vowel spreads, just as in compensatory lengthening (I only draw the three relevant segments of *ku n daba*:

```
                    x  x  x
                    |  |  |
(117) a. Input:  k  u  n  d  aba

                     x  x  x
                     |/ /|
      b. Output: k   u  n  d  aba
```

Other candidates for representations with two segments being linked to one timing slot are affricates (e.g. [c] = /t/+/s/ linked to one slot) and doubly articulated consonants (e.g [$\widehat{\text{kp}}$]).

4.4 Feature Geometry

The Place node

If features are organized into tiers, we still have to find out how those tiers are related to each other. We have now established that there is evidence for one central timing tier, the skeleton. But this still leaves many different options. One possibility — maybe the simplest one — is to assume that all features are linked directly to this one central tier. This is sometimes called the *bottle brush* model:

```
                [Labial]
                   |
[Coronal]  —   x  —  [Dorsal]
                   |
                [Voice]
```

(118)

However, there is evidence against this simple model, and pointing in the direction of features being organized in arborescent structures; the school of thought is called *feature geometry* (using a somewhat excentric definition of the term 'geometry'). The most straightforward evidence here comes from the fact that sometimes certain features group together. A well known case is *place assimilation*. In many languages of the world, nasal consonants assimilate in place of articulation to the following consonant. The following examples are from Chuckchi (Palaeosiberian, Siberia), where we assume the assimilating nasal is ŋ underlyingly.

feature geometry

place assimilation

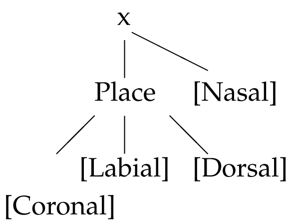
(119)	təŋ-əɬ-ən	'good'
	tam-pera-k	'to look good'
	tam-vairgin	'good being'
	tam-waɤərɤ-ən	'good life'
	tan-t^{s}ai	'good tea'
	ten-leut	'good head'
	tan-ran	'good house'
	ten-yəɬqət-ək	'to sleep well'

In these cases, the nasal assimilates in the value for the features [Coronal], [Dorsal] and [Labial], but not for any other feature (e.g. it does not lose its nasality or turn into a fricative).

We could of course assume that Chuckchi has three different phonological rules which we could informally state as follows:

(120)
a. Spread [Coronal] from a consonant to a preceding nasal.
b. Spread [Dorsal] from a consonant to a preceding nasal.
c. Spread [Labial] from a consonant to a preceding nasal.

But this is very unattractive, especially because we find a similar phenomenon in many languages of the world, and it always involves these features. But more in general, we would want to give a uniform description of phenomena such as this. In order to achieve this, we posit an organizing node in our phonological representations, called a Place node. The place nodes are not linked individually to the central skeleton, but through this organizing node:

(121) 

We can now formulate the relevant rule in a very simple and straightforward way:

(122) Spread the Place node from a consonant to a preceding nasal.

When we spread the place node, we spread all the relevant features at the same time. Nasal assimilation thus gets a simple and straightforward formalisation.

Another type of evidence pointing in the same direction comes from *debuccalization*. For instance in certain dialects of Malay, consonants in coda position change according to the schedule in (123) (Humbert, 1995; Botma, 2004):

debuccalization

(123) a. /p, t, k/ → [ʔ]
b. /s, f, h/ → [h]
c. /m, n, ŋ/ → [N] (a 'placeless nasal')

(124) a. /ikat/ → [ikaʔ] 'to tie'
b. /lipas/ → [lipah]
c. /ʔawan/ → [ʔawaN]

The traditional name for this process is 'debuccalisation', since all the oral articulators become inactive. On the other hand, the manner of articulation stays constant: a stop /t/ stays a glottal stop [ʔ], a fricative /s/ stays a fricative /h/, and a nasal /n/ stays a nasal, albeit a placeless one.

Again, we could formulate this in terms of three independent rules:

(125) a. Delink [Coronal] at the end of the syllable.
b. Delink [Labial] at the end of the syllable.
c. Delink [Dorsal] at the end of the syllable.

This would come at a loss of generality, however, especially since again the three processes seem often linked. For instance, the same phenomenon can be found in London English (Lass, 1976; Gussenhoven and Jacobs, 1998).

(126)

piʔtʃə	*picture*
mɪiʔ wɪljəm	*meet William*
nɪʔ næʔs	*knick-nacks*
nɔʔ næʊ	*not now*
kɪiʔ smɑɪlɪŋ	*keep smiling*

Introducing a Place node allows us to simplify the formalism considerably. Both Malay and London English are subject to the following rule:

(127) Delink the Place node at the end of the syllable.

Note that this means that we assume that segments such as [ʔ] and [h] lack a place of articulation node. It is not the case that these segments have a specification [-coronal, -labial, -velar]: they do not have any place features whatsoever.

This particular assumption also makes it easier to understand why the glottal stop very often functions as the 'default consonant'. For instance, we fill in this consonant in German if otherwise a situation of **hiatus** – two adjacent vowels – would ensue, or if a word starts with an open syllable:

(128) *Theater* 'theatre' [teʔátʁ], *Chaos* 'chaos' [káʔɔs], *atmen* 'to breathe' [ʔátmən]

The reason why a consonant has to be inserted here, probably is the same as why we have liaison in French:

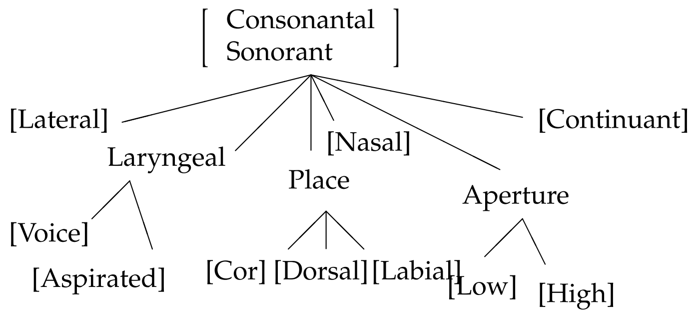
(129) An x-slot has to precede the vowel in a syllable.

Different from the liaison context, there is no obvious neighbouring consonant to fill the empty slot in cases such as in (128). Therefore the slot is filled by the phonological rule component. We can understand why it is the glottal stop that is inserted in contexts like this, if we assume some principle of representational economy: if we have to insert something, we prefer to insert as little as possible to satisfy our needs. If we need to insert a consonant, it is better to insert one where we do not have to include a Place node (and Place features).

It is not the case, by the way, that glottal stop is the default consonant in all languages of the world. Some languages do not allow this type of segment at all — apparently, they disfavour Place-less consonants. In such cases, some other consonant such as /t/ fulfills that role.

The Feature Tree

The next question obviously is whether the Place features would be the only ones which are organized into a separate node. Most phonologists in the feature geometry paradigm would agree that this is not the case, and that there is more internal organisation to the segment. Although there is no general agreement on this point, the following structure may be considered as fairly representative for the mainstream:



(130)

Further structure is possible; for instance, Place and Aperture are often combined into a Supralaryngeal node, combining all the instructions for organs above the larynx. Also, the position of the features [Continuant], [Nasal] and [Lateral] has been the topic of debate.

It needs to be observed that the claim underlying virtually all work in Feature Geometry is that the structure in (130) — or whatever should be replacing it — is universal: if a language has a feature [Continuant], it will be organized into the structure as indicated.

Nodes beyond Place

A prediction of this model is that all the organizing nodes should behave like the Place node. There should be processes — for instance of assimilation — which involve exactly the features that are dominated by some node and none of the others. We will briefly review some of this evidence for the Aperture node and the Laryngeal node.

As to the former, consider the following examples from Brazilian Portuguese:

(131)

2nd person			1st person		
/mɔr-a-s/	[mɔ́ras]	'you reside'	/mɔr-a-o/	[mɔ́ro]	'you reside'
/mɔv-e-s/	[mɔ́ves]	'you move'	/mɔv-e-o/	[móvo]	'you move'
/sɛrv-i-s/	[sɛ́rves]	'you serve'	/sɛrv-i-o/	[sírvo]	'you serve'

Like in many (Romance) languages, verbs in Portuguese have a so-called theme vowel, which behaves in some respects like a suffix, but which at the same time is determined by the stem: the verb 'to reside' has -/a/- as its theme vowel, 'to move' has -/e/-, and 'to serve' -/i/-. These theme vowel surfaces for instance in the second person singular, which has the consonant-initial suffix /s/, as is illustrated in the lefthand column. However, the first person singular suffix is -/o/, and this may be a reason why the theme vowel disappears — otherwise we would again create a hiatus.

But when the theme vowel disappears, something happens to the stem vowel: it changes from /ɔ/ to [o] in 'to move' and from /ɛ/ to [i] in 'to serve'. These are changes in vocalic aperture: /ɔ, ɛ, a/ are low vowels ([+low,-high]), /e,o/ are mid vowels ([-low,-high]) and /i/ is a high vowel ([+high,-low]). What happens, then, is that the stem vowel takes over the aperture features of the disappearing theme vowel. In autosegmental terms, we can describe this as relinking of the Aperture node, rather than the individual relinking of the features [±high] and [±low].

The argument for the Aperture node thus comes from relinking; we will provide an argument in favour of the Laryngeal node from neutralisation. Korean has three series of stops, traditionally called voiceless, 'tensed' and aspirated (Rhee, 2002). There is no general agreement as to what exactly are the phonetic or phonological correlates of these three dimensions, but it is clear that they have to be described by Laryngeal features. It is also clear that they can contrast in a position before a vowel:

(132)

lenis	fortis	aspirated
[pal] 'foot'	[p'allɛ] 'laundry'	[pʰal] 'arm'
[tal] 'moon'	[t'al] 'daughter'	[tʰal] 'mask'
[kɨn] 'root'	[k'ɨn] 'string'	[kʰɨn] 'big'

However, at the end of the syllable, we only find the lenis variants:

(133)

lenis	fortis	aspirated
[cip-to] 'hous EMPHATIC'	*[cip']	*[cipʰ]
[mit-to] 'bottom side EMPHATIC'	*[mit']	*[mitʰ]
[puək-to] 'kitchen EMPHATIC'	*[puək']	*[puəkʰ]

This looks very similar to a process which we know from languages such as Dutch, German, Turkish and Catalan and which is usually called *final devoicing* (the example is from Dutch, in case anybody did not realize):

(134) a. Beginning of syllable:

voiced	voiceless
[dɑk] 'roof'	[tɑk] 'branch'
[bɑk] 'bin'	[pɑk] 'suit'

b. End of syllable:

voiced	voiceless
*[hɔnd]	[hɔnt] 'dog'
*[ɛb]	[ɛp] 'ebb'

For Dutch — as well as the other languages just mentioned — it may be assumed that what is going on is that the feature [+voice] gets lost at the end of the syllable; the remaining structure is then interpreted as voiceless. Korean shows the same phenomenon, but with one difference: at least two different features have to be lost — the ones distinguishing tensed and aspirated consonants from lenis ones. Again, this can be profitably described if we assume that the relevant rule is something like the following:

(135) Delink the Laryngeal node from a consonant at the end of the syllable.

This rule can even be applicable to the final devoicing languages such as Dutch; in these languages there is only one Laryngeal feature, so it is hard to tell *a priori* whether it is just this feature which is delinked, or the node dominating it.

Root nodes and skeletal points

There is one more organizing node to be discussed: the *root node*, the node to which all other organizing nodes, as well as individual features, are eventually attached. This is the node in (130) which carries the features [Consonantal, Sonorant].

The fact that the root node carries these features has an important implication under autosegmental assumptions: we cannot spread either one of those features independently, they will always carry the other features along, because they are attached with them. Whereas it is possible to spread e.g. [Nasal] without spreading any other part of the tree, spreading of e.g. [Consonantal] will always result in *total assimilation*, a famous instance of which is found in the Lesbian and Thessalian dialects of Ancient Greek, where /s/ assimilated completely to an adjacent sonorant segment:

total assimilation

(136)	*g^{w}olsā	>	bollā	'council'
	*awsōs	>	awwōs	'dawn'
	*esmi	>	emmi	'I am'
	*naswos	>	nawwos	'temple'

(Notice by the way that again we are not dealing with a synchronic phonological rule in this case, but with a phonological change; which is not necessarily the same thing.)

What is impossible, according to this model, is a change where a sonorant would change to a stop with exactly the same place features due to assimilation:

(137) amta > apta (impossible change; and impossible phonological rule).

Another implication of these assumptions, and of the analysis underlying (136), is that the root node organizes all the features, but is still distinct from an x-slot; for we see the process happening in (136) as spreading of the root node with all its features from one x-slot to the next.

This assumption seems necessary also for most of the analyses we presented above, where it was equally the case that all the features spread together from one skeletal point to the next.

At the same time it may be seen as a little unfortunate that we now have two tiers which organize all the segments. Furthermore, there is an empirical problem with this particular implementation of segmental structure in autosegmental phonology. We know that complex segments can be for instance *affricates* (sharing place features but differing on continuancy: [t͡s, p͡f]), prenasalised segments (sharing place features but differing in nasality: [ⁿd, ᵐb]), or doubly articulated stops (sharing all features except for place). There has never been found any evidence for complex segments where the two parts differ on many different dimensions (e.g. *[t͡ŋ], *[p͡a]). This is unexpected, given the autosegmental model.

affricates

As a methodological aside, note that an assumption underlying this criticism is that every structure which can be generated by the formal model, also needs to be attested in some of the world's languages. In principle, it is of course possible that structures such as *[p͡a] do indeed exist, but only in languages which have not yet been considered in sufficient detail: we simply do not know about them yet. However, it is good practice in phonological theorizing to assume that structures do not exist until somebody points out that we do need them in the analysis of some language. If we would not take this as our guideline, it would be almost impossible to compare theories: a model which would say that 'anything goes in natural language' would beat everybody else; but it would not be very interesting. In other words, we try to make our model as *restrictive* as possible. The model developed so far is not **restrictive** restrictive enough from this point of view; it *overgenerates*. **overgenerates**

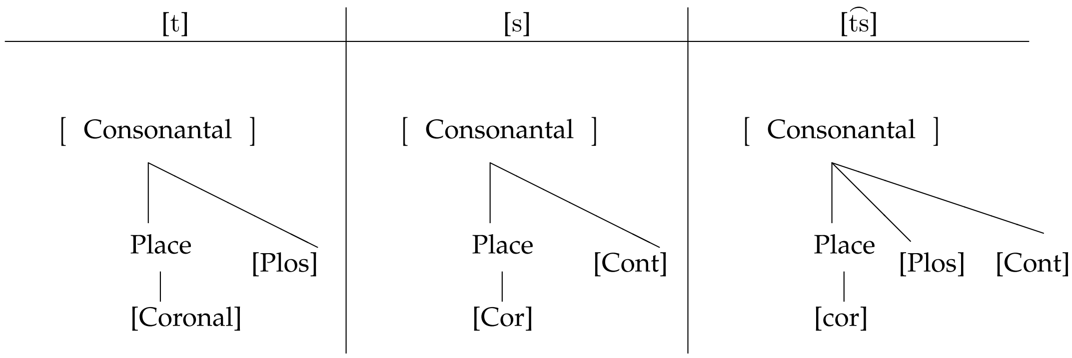
This problem still awaits a full formal solution at present. Somehow we have to assume (without an explanation) that one timing slot cannot host more than one root node. Therefore, we have to find a different representation for complex segments.

From what we have seen so far, we can already conclude the following:

(138) Complex segments bear more than one feature (value) of a specific type.

For instance, [t͡s] is exactly like [t] and [s], except for one point: whereas [t] is [+continuant] and [s] is [-continuant], [t͡s] is *both* [+continuant] and [-continuant]. Heavily simplifying our feature trees, we can draw the three segments as follows:

(139)

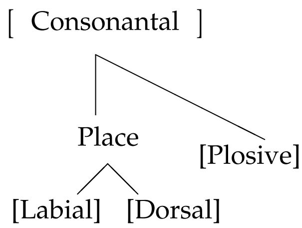


In the structure for the affricate, two features (on the same tier) are now linked to the same segment (in this particular case, to the same root node). This parallels two tones being linked to one segment. A similar picture can be drawn for prenasalised segments.

As an aside, note that such representations seem to be slightly easier to argue for in terms of binary features, rather than the unary features we have adopted here: rather than having to adopt features [Plosive] and [Continuant] or [Nasal] and [Oral], we can say that these are two instances of the same feature, [±Continuant] or [±Nasal].

Multiply articulated segments might be a little bit different; the following represents [k͡p] (again, abstracting away from certain complexities):

(140)



This representation is different because the two Place features are probably not on the same tier; they are linked to the same node, but they still represent different dimensions. Because they are not on the same tier, they are also not temporally ordered with respect to each other; which gives the (correct) prediction that they are realised at the same time. In order to produce a labiovelar consonant, for instance, one has to put the lips in the position of /p/ and the tongue in the position of /k/ before the release that is typical of plosives.

4.5 Possible changes and markedness

We are now in a position to refine our view of underlying representations and the computational mechanisms needed to change from one to the other, layed out in Chapter 3.

It is useful to have a list of possible phonological changes that an underlying representation can undergo from the underlying form on its way to the surface. Now that we have a more refined view of how features relate to each other, we can improve our view of what is a plausible change from an underlying representation to a surface representation.

We have already seen that such changes often seem to be synchronic reflexes of diachronic change. A morpheme changes its shape in some circumstances, but not in others. We then suppose that this has the effect that a synchronic change is applied to the underlying form in the phonology. This means that the language user apparently has this computational power at their disposal: we do not just store forms, but we can *do* things with them after having applied morphosyntax.

markedness

The reason why changes are applied is very often one of so-called *markedness*: the resulting form is 'less marked'. This term is very important in grammatical theory, although it has appeared to be remarkably difficult to define. It refers to some formal simplicity (a segment with three features is less marked than a segment with four features), but also to typology (a segment that occurs in more languages is less marked than a segment that occurs in fewer languages) and to frequency (unmarked segments are more frequent in a given language than marked ones). It sometimes is also meant to refer to learning: a segment that is learned earlier is less marked. And finally it refers to language-internal frequency: a segment that occurs more often is less marked.

There is a claim that these different criteria for markedness converge, and in a surprising number of cases, that claim seems to hold; but this is not always the case, and it is important to keep this in mind. Above I have ordered the criteria in their order of importance for theorizing.

Feature changes

Several of the examples we have seen in this chapter (final devoicing in German, palatalisation in Southern Kongo) involve feature changes. We have already seen several possible changes in chapter 4 on Autosegmental Representations. As a matter of fact, most arguments in favour of autosegmentalism derive from alternations: it is the way in which phonological representations behave when they get in touch with other representations.

First, features can be **delinked**. Final devoicing might be an example: the word *Hund* ends in a *d* that is carrying the feature [Voice], that feature gets delinked to get the output form [hunt]. It then becomes floating, or it is even completely deleted from the representation.

The reason for delinking is very obviously one of markedness according to the above definition: the representation of [t] in German is simpler than that of [d]. Markedness can also be positionally defined. We will see in chapter 5 that the end of the syllable is a 'weak' position, less likely to sustain marked (i.e. complex) segments.

At the same time, the fact that devoicing is a common process at the end of syllables, but e.g. loss of place (although that happens in some languages as well) is less so, most probably does not have a formal reason, but should rather be explained phonetically. Voicing is particularly difficult to observe in obstruents at the end of a syllable; more so than place of articulation. So the diachronic change underlying the synchronic process is more likely to occur for voicing than for place.

Delinking can also be a response to the OCP. In this case, we also talk of *dissimilation*. Here is an example from Ancient Greek, in which the underlying fricative /θ/ of the passive suffix turns into a plosive [t] if it occurs next to a stem final fricative:

(141) a. agap-i-θik-e 'he was loved'
fer-θik-e 'he was carried'
stal-θik-e 'he was sent'

b. akus-tik-e 'he was heard'
ðex-tik-e 'it was received'
ɣraf-tik-e 'it was written'

We can clearly view this as a case of OCP on the feature [Continuant] in a strictly adjacent position (as the word for 'he was carried' is not affected'). As a response, one of the two features gets delinked.

Another possible feature change is **spreading**. We have seen many instances of this in chapter 4, such as place assimilation in English nasals, or palatalisation (if we assume this is spreading of a feature of the front vowel to the preceding consonant). If a feature spreads, the target segment obviously does not become less marked, since it even aquires a feature. From the point of view of representation even the word as a whole does not become less marked, since we end up with the same number of features, one of them which is now linked to more than one segment. Yet from a phonetic point of view, spreading always leads to less markedness, since some articulatory gesture can be sustained for a longer period.

More problematic for a simple definition of markedness is the existence of **feature addition**. Examples of this are rare and may often be subject to reanalysis (the phenomenon which seems to argue in favour of feature addition can also be understood in a different way). A possible class of 'real' examples is so-called *phonetic enhancement*. Sometimes a phonological contrast in a language is expressed along several different phonetic cues. For instance, in a five vowel system [i, u, e, o, a], the vowels [i] and [u] are distinguished both in the frontness-backness dimension and in the the roundedness dimension, at least under the assumptions of feature theory, and similarly for the [e] and [o]. We thus make a necessarily arbitrary choice which of the two contrasts is the 'real' one; but no matter what we do, at some point we have to add the other feature to our representation to make sure that the phonetics gets both right.

phonetic enhancement

The term 'enhancement' gives us a nice description of why this happens: the contrast becomes easier to perceive. Rounding and backness have a similar effect on the acoustic signal since they both make the resonance chamber for the vowel longer, rounding by protuding the lips and backness by placing the tongue at the end of the oral cavity. It is thus convenient to have both features available.

Enhancement runs against reduction of markedness — they are different forces, but they also seem to operate at different levels. Markedness is a force within the phonology of human languages, whereas enhancement seems to operate mostly at the interface with phonetics, after most phonological computation has been done. There are no known examples of features, one enhancing the other, which can for instance be shown to spread independently from each other.

There are also a few cases where it seems that a feature 'moves' from one segment onto another. We have seen examples of this in chapter 4, and we have also seen that autosegmental phonology allows us to decompose such 'movement' as essentially a combination of spreading and deletion (or delinking).

Segmental changes

Another group of changes are segmental. We distinguish again between three subgroups:

The first of these is **deletion**, again a straightforward case of a reduction of markedness. We have seen an example above in the Tibetan numeral system, where the first consonant of an underlying word-initial cluster is deleted. In this, like in mamy other cases of segmental change, the reason is most probably a so-called phonotactic one: words in a language tend to arrange vowels and consonants in certain ways and not others. Clusters of consonants are often avoided, and deletion of one of those consonants is a method for achieving this.

Another method is insertion of segments, also called **epenthesis**. For instance in Mongolian (Mongolic, Mongolia), the word for 'filler' underlyingly is /šitms/. The final cluster of that word is broken up by a schwa on the surface: [šitəms]. Such a structure may be less marked if we just count features, because the schwa is obviously adding feature material to the overall representation, but instead of one very complex syllable we now have two much less complex syllables and as we will see in Chapter 5, that this counts as less marked.

Furthermore, typically epenthetic vowels tend to be the least marked vowels in the vowel inventory of the language. If a language has an underlying schwa, like Mongolian, the epenthetic vowel tends to be schwa as well. We need to insert a vowel, but we still want to make the overall representation as little marked as possible.

Another possibility is that features of surrounding segments are spread to the epenthetic segment. When the epenthetic segment is adjacent to the source of those features, we may call this *lengthening*. In some other cases, the source segment is a few segments removed; this is more common for vowels than for consonants. We then have an epenthetic vowel with the same content as a vowel in an adject syllable. This we can call *copying*. This is another way of inserting a segment without raising overall markedness too much, since all of the features are already present in the underlying representation.

lengthening

copying

The last possible segmental change is **moving its position**. This process is called metathesis, and its existence as a synchronic process is not entirely uncontroversial, but a fairly strong example can be found in Leti (Autronesian, Indonesia), in which certain words have a different order of consonants and vowels in the middle of a phrase than at the end of a phrase:

(142)

Phrase-Final	Phrase-Medial
hline *urnu*	*urun moa*
'breadfruit'	'Moanese breadfruit'
bubru	(bubur vetra)
'porridge'	'maize porridge'
βura	*βuar lavan*
'mountain'	'big mountain'

As you can see the last vowel and consonant of the word switch position depending on the syntactic position. Exploring the precise reason why this happens, would lead us a little bit too far afield, but it clearly is again some form of phonotactics, albeit now within a larger domain than the word.

In the interaction with morphosyntax, we can see a few more processes, like reduplication in which it seems that a segment is *copied*. Such processes will be discussed in Chapter 8, where we will show how they can be decomposed into insertion of segmental positions and autosegmental spreading of feature material.

Within phonology itself we can find another processes that can also be decomposed into the smaller steps we have seen above. One of them is *coalescence* in which it looks as if two segments become one. Here is an example from Indonesian (Austronesian, Indonesia):

coalescence

(143)
a. /məŋ/ + /ikat/ → [məŋikat]
b. /məŋ/ + /pilih/ → [məmilih]
c. /məŋ/ + /tulis/ → [mənulis]
d. /məŋ/ + /kasih/ → [məŋasih]

The prefix /məŋ/ can be combined with a noun to form a verb. When the original noun starts with a consonant, the cluster of the prefix final /ŋ/ and the stem-initial consonant is replaced with one consonant which seems to inherit the nasality of the /ŋ/ and the place of articulation of the stem consonant: the two have 'coalesced' into one with properties of both.

Such a process can be understood as a combination of spreading of the place features from the plosive to the nasal (i.e. place assimilation) and deletion of one of the consonants in a cluster (for phonotactic reasons). Both are attested independently in languages of the world; in Indonesian they happen to be combined, as we have seen.

This thus concludes our overview of possible changes to phonological form: the six processes mentioned in this section — delinking, spreading and insertion of features, deletion, epenthesis and possibly moving of segments — basically are what phonology can actively do to underlying forms.

4.6 Exercises

1. Give the tonal representation of 'they looked' and 'we sent them' in Kikuyu, taking (70) as your model.
2. In Mende, vowels in monosyllabic words can have one of five tones: high, low, rising, falling, or first rising and then falling. In words with two syllables, the first syllable is always high or low, and the second syllable is high, low, or falling; a falling tone only occurs after a low toned first syllable. Finally in words of three syllables, all syllables have only a high or a low tone. How can you explain this pattern autosegmentally? (You can make up your own examples if you want to illustrate.)
3. Look at the following forms in Chizigula, in which high toned syllables have an accent and toneless syllables do not (*ku* is a prefix):

   (144) *kudamanj* 'to do', *kudamanjiza* 'to do for', , *kudamanjizana* 'to do for each other', *kulombéza* 'to ask', *kulombezezána* 'to ask for each other', *kulombezéza* 'to ask for'

4. In a secret language in Thai, words are changed a little bit so that outsiders cannot understand them: *khluâì hòóm* 'banana' is pronounced as *khlóòm huàí* and *téǹ ram* 'dance' as *táḿ ren* (acute accent denotes a high tone, grave accent a low tone, no accent on a vowel is a mid tone). How can this be constructed as an argument for autosegmental phonology?
5. In some varieties of Latin American Spanish something remarkable happened to the [s] at the end of a syllable. Give a precise description of what has happened in autosegmental terms (the circle under a vowel indicates that the vowel is voiceless):

(145)	European Spanish	L.A. Spanish	
	mismo	mii̥mo *of* mim̥mo	'same'
	fosforo	foo̥foro	'match'

6. In an innovative variety of the dialect of Shanghai we find an interesting tonal pattern. Consider first the following underlying representations for several morphemes (tonal specifications are added in parentheses; M denotes a mid tone, a third tonal level in some languages):

- *çi* 'fresh' (HL); *wa* 'yellow' (LH); *du* 'big' (LH) ŋ 'fish' (LH); *ço* 'small' (MH)

Now consider the tones of following phrases:

- ço+ŋ → ço (M) ŋ (H) 'small fish'
- çi+ŋ → çi (H) ŋ (L) 'fresh fish'
- wa+ŋ → wa (L) ŋ (H) 'yellow fish'
- ço+wa+ŋ → ço (M) wa (H) ŋ (L) 'small yellow fish'
- çi+wa+ŋ → çi (H) wa (L) ŋ (L) 'fresh yellow fish'

Describe what is going on here in autosegmental terms, and give the tonal pattern for 'big yellow fish'.

7. Consider the following forms in Tiberian Hebrew.

(146) a. *seefer* 'book'
b. *geʃem* 'rain'
c. ʔiːʃ 'man'
d. *haːr* 'mountain'

Now look at the following forms, to which we added the definite determiner *ha*:

(147) a. *hasːeefer* 'book'
b. *hagːeʃem* 'rain'
c. *haː*ʔiːʃ 'man'
d. *haːhaːr* 'mountain'

Assume laryngeals and pharuungeals cannot geminate. Give a description of what happens.

Discussion and further reading

The Kikuyu data are wonderfull described and analysed in Goldsmith (1990);so are the Luganda data later on in this chapter.

The Margi data are from (Hoffmann, 1963; Williams, 1976; Kenstowicz, 1994)

Since the work of Odden (1986), it is no longer assumed that the OCP is a universal principle, but it can still be seen at work as a tendency in some languages.

The Shona data are discussed in Odden (1980); Myers (1987); Kenstowicz (1994)

The Bergen Dutch data are from van Oostendorp (1998)

The interaction between Rendaku and Lyman's Law is the topic of extensive literature. See for example Itô and Mester (2003).

The discussion of Finnish follows Keyser and Kiparsky (1984); Gussman (2002). Compensatory lengthening in Old English was described in (Ewen and van der Hulst, 2001; Gussman, 2002); the Turkish data have also been amply discussed, for instance in Sezer (1986); Goldsmith (1990); Kenstowicz (1994); Gussman (2002). Italian Raddioppiamento Sintattico has been described by (Nespor and Vogel, 1986).

The Chuckchi data are in Odden (1987); Clements and Hume (1995).

Syllables

5.1 Evidence for syllable structure

Where in the world can we find syllables? The best place to look for relatively straightforward evidence might be in poetry. In many poetic traditions, every line in' a poem has a fixed number of syllables. Very famous in this respect is the Southern Slavic epic tradition: many poems in Serbian and Croatian poetry consist of lines of exactly ten syllables each (*decasyllables*; *deka* is the Greek word for *ten*): **decasyllables**

(148) *Što se bili u gori zelenoj?*
What itself be-white in mountain green
al su snizi, al su labutovi?
or is snow, or is swans
da su snizi, već bi okopnili,
if were snow, already it melted
labutovi već bi poletili
swans already fly-away
"What is white on the green mountain? Snow or swans? If it were snow, it would already have melted away; and swans would already have flown."

These are the first lines of 'The Mourning Song of the Noble Wife of the Asan Aga' (Asanaginica), a folk ballad from 1646-49 and a region which currently belongs to Croatia. Although decasyllables are still occasionally composed by modern authors in the region, their origin lies in a medieval oral tradition; they were composed by poets who did not necessarily *write* their poems; as a matter of fact, authors might have not been able to write. This, and the fact that there are other traditions around the world which 'count syllables' in poetry, shows that the syllable is an intuitive concept for human beings, and does not necessarily depend on their literacy.

Similarly, we can observe that people tend to find it easier to count the number of syllables in the word *syllabicity* than the number of segments in *sounds*. (You can do a small experiment with people around you, asking them e.g. how many syllables *encyclopedia* has, and how many vowels and consonants *journal* has. The answer is the same, but the first is easier to come by.) There is evidence that children know how to syllabify words before they can divide them into segments and before they can write.

At the same time, it turns out rather difficult to give a precise definition of the notion 'syllable' that covers all the intuitions. In addition, syllables so far have n not been straightforwardly detectable in the acoustic signal, possibly even less so than individual segments. Even so, the phonological evidence in favour of the syllable is manifold. Many different observations can be phrased in a much more elegant way if we accept the notion of a syllable. Furthermore, the syllable comes very close to being universal: most, if not all, languages give evidence for it.

We will now review the most important types of evidence that have been put forward for recognizing the syllable as a formal unit of phonological analysis. We will see that this evidence comes from many different angles, and converges on two properties that syllables have:

constituents

- They are *constituents*, i.e. groups of smaller units (segments, in this case) which behave as a unit in some ways.

headed

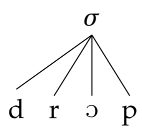
head

dependents

- They are *headed*, i.e. one of the smaller units is more prominent than the others and determines the properties of the constituent as a whole. We call this prominent unit the *head*; the other units are called the *dependents*. In syllable structure, the head is typically a vowel.

We will discuss this terminology in more detail in section (5.2). For now, we should note that headed constituents are also known in other branches of grammatical description, such as morphosyntax.

In order to represent these two properties, phonologists draw diagrams like the following (for the English word *drop*); we have already introduced such representations before:

(149) 

See Section ??

The Greek letter σ (*sigma*) denotes the syllable node, and the lines — which are sometimes thought of as association lines — the fact that segments belong to the constituent. The vertical line between σ and ɔ is special and denotes the relationship between the head and the constituent. A more refined view of this structure will be given in section 5.2.

Stress

The most important argument in favour of syllables is consistent asymmetries between vowels (V) and consonants (C) in the way they are organized into words. There are many phenomena where the vowels can 'see' each other at a distance across consonants, whereas there are very few cases where consonants could see each other across vowels.

word stress

See Chapter 7

One such phenomenon long distance contact between vowels is *word stress* . Many (albeit not all) languages have one vowel in the word standing out for being more prominent than the others: it has a higher pitch and/or it is longer and/or it is louder. A simple example is Southern Peruvian Quechua. Stress in this language is on the penultimate vowel (with some exceptions). Stress is indicated by an accent ˈ before the vowel that is stressed.

(150) a. wˈassi "house"
b. wassˈipi "in (the) house"
c. wassikˈuna "houses"
d. wassikunˈapi "in (the) houses"

In order to determine where the stress falls in these words, one needs to count vowels. The absolute number of segments is irrelevant: in (150a) the stressed vowel is the fourth segment counted from the end, whereas in the other words, it is the third segment. In order to determine the exact position of the stress, we thus should not count segments, but we should count vowels from the end and disregard the consonants.

Importantly, there are no similar known processes which apply to the second consonant from the end and disregard all the vowels. We can account for this in the following way: we assume that stress is not so much a property of vowels as it is of syllables. We thus assign stress to the second *syllable* from the right. However, we also assume that this property is first and foremost expressed on the head of the syllable, which is typically the vowel. Since consonants are not typically heads, but dependents of syllables, it is not possible to apply a similar technique to them to get a system to count consonants and not vowels.

An even stronger argument comes from those languages in which not just the distance from the word edge plays a role in terms of stress but also the structure of the individual syllables. Many dialects of Arabic (Afroasiatic) are of this type. Let us briefly consider Palestinian Arabic as an example. In this language, stress falls preferably on the rightmost *heavy syllable* (i.e. a syllable that is closed by a consonant or that contains a long vowel; as opposed to a *light syllable*, which ends in a short vowel). If there is no such syllable, then stress falls on the first syllable of the word (the final syllable does not count): **heavy syllable** **light syllable**

(151) a. *Words with a heavy syllable*:
i. [baˈʃuːfiʃ] 'I don't see' ([ʃuː] has a long vowel, and hence is heavy)
ii. [kaˈtabti] 'you FEM SG wrote' ([tab] is closed and hence heavy)
b. *Words without a heavy syllable*
i. [ˈdarabu] 'they hit'
ii. [ˈzalama] 'man'

It would be difficult to describe this system without referring to syllables and their structure, but just considering vowels and consonants, as you can try out for yourself. Furthermore, there are many languages in the world that make this distinction between heavy and light syllables, as we will see in Chapter 7. This therefore gives us a piece of evidence that is not direct (we establish the existence of the syllable not by directly observing that it is there) but still very forceful, in particular because it turns out that we need the concept to understand many other, at first sight unrelated phenomena.

Reduplication

Another phenomenon where we find evidence for the syllable is in *reduplication*, a morphological process in which a part of a stem is copied to express a **reduplication** certain meaning. In almost all known cases, the copied part corresponds to a phonological constituent; in several languages this is a syllable.

Yaqui (Uto-Aztecan), for instance, reduplicates the first syllable of the word (there is a lot of discussion in the literature as to what the semantics of this reduplication really is; it seems to be a more intense reading of the verb stem):

(152)	a.	vu.sa	vu.vu.sa	'awaken'
	b.	chi.ke	chi.chi.ke	'comb one's hair'
	c.	he.wi.te	he.he.wi.te	'agree'
	d.	ko.a.rek	ko.ko.a.rek	'wear a skirt'
	e.	vam.se	vam.vam.se	'hurry'
	f.	chep.ta	chep.chep.ta	'jump over'
	g.	chuk.ta	chuk.chuk.ta	'cut with a knife or saw'
	h.	bwal.ko.te	bwal.bwal.ko.te	'soften, smooth'

The dot between two segments denotes the syllable boundary. The examples in (152a-d) show that if the stem starts with an open syllable, a sequence of a consonant followed by a vowel ('CV') is copied, whereas (152e-h) show examples of a stem starting with a closed syllable, which is also faithfully copied as a consonant-vowel-consonant sequence ('CVC'). Describing such a morphological process requires referring to the syllable, and in this sense, then, reduplication provides another piece of indirect evidence for its existence. Yaqui children learn how to form this intensifier construction, which means that they can have no problem operating syllables.

Language games

Language games also sometimes refer to the concept of the syllable. A case in point is *Vesre*, a secret language originating in the underworld of Argentina (around Buenos Aires) and Uruguay, and sometimes used in tango lyrics (for instance in the famous tango song "¿Qué querés con ese loro?"). *Vesre* takes a Spanish word and puts its syllables in the inverse order:

(153)

	Spanish	*Vesre*	gloss
a.	pizza	zapi	'pizza'
b.	caballo	llobaca	'horse'
c.	réves	vesre	'inverse'

Other languages have similar procedures in language games (e.g. French *Verlan* and Tagalog *Binaliktad*), showing that this is a procedure that apparently shows up with some frequency also in types of informal language, and is apparently used there without problems by all kinds of speakers.

Another common type of language game is one in which a syllable is added before or after every syllable of the original word, as in the following example from Hausa (Chadic) in which *da* is prefixed to every syllable of the word:

(154) tsíntsíyáa → dá-tsín-dà-tsìi-dà-yáa 'broom'

(Notice that interesting things also happen to the tones, but we will ignore this here.) In order to be able to play this game, one needs to be able to divide the word into syllables in the first place; further, it probably is no coincidence that the prefixed sequence is also a syllable. (At this point we obviously do not yet have a precise definition of where the syllable boundaries are; we will return to this later in this chapter.) See Section 5.2

The crucial aspect of such language games is that they play a role in a context that is playful and oral. Just like in the case of the South Slavic poetry at the beginning of this chapter, this demonstrates that ordinary language users do not find it difficult and do not need a high level of literacy for ordinary language users to acquire such a system, or to understand it. This is not true for sequences that form phonological non-constituents. There are apparently no language games in which one would take random segment sequences and revert them or prefix them by other material.

We should take approach from language games with some caution, however. There is obviously always something 'artificial' about them: the rules of these games have been made up with the purpose of playfulness or deception and there is no 'real' morphological process which shifts around syllables in this way. Nevertheless, it is suggestive that human creativity in inventing succesful language games seem to have boundaries, and these boundaries are among other things defined by constituency.

Psycholinguistic evidence

Another type of evidence concerns several type of data from psycholinguistic research. One type concerns speech errors. For many decades, psycholinguists have been studying the mistakes which people occasionally make from a phonological point of view. One of those is so-called *blending*, in which two words get inadvertently mixed. An example is *yout*, which blends the beginning of *yell* and the end *shout*. **blending**

We can call the point where one word turns into the other (in this case, this happens between *y* and *out*) the *break*. Where do these breaks occur exactly? Already in 1972, the psychologist D.G. McKay studied this question, based on a corpus of mistakes made by university professors in Vienna at the end of the 19th Century! McKay did something simple with these very old data: he looked whether breaks occurred within syllables (such as in the case of *yout*) or at the boundary between two syllables (e.g. in *war.der* from *warmer* and *col.der*). Here are his results: **break**

(155)

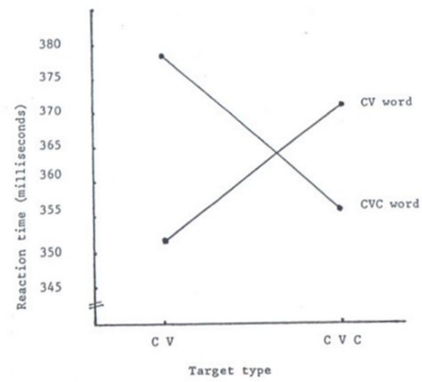
	Breaks within syllables	Breaks between syllables
Data%	40	60
Chance %	64	36

The 'chance' level means that a clear majority of 64% of all boundaries between segments are within syllables (which means that syllables span more than two segments on average). Still, breaks are much more likely to occur at the boundaries between two syllables than between any two segments within a syllable. This shows then that somehow syllables play an important role in language planning, and this has been confirmed by many studies since, also for speakers who were not Viennese university professors.

There is also evidence that the syllable plays a role in the *perception* of speech. In a famous experiment, a group of French speakers listened to long lists of individual words. They were asked to find a specific sequence in these words, for instance *pa* or *pal*, and press a button if they heard this sequence. The trick was that the word list would contain items that started with the same three segments (in our example *p-a-l*), but which were syllabified in a different way; in our case these would be for instance *palace* 'palace' which has an initial syllable *pa*, and *palmier* 'palm', which has an initial syllable *pal*. Other pairs were *carotte-carton, tarif-tartine, garage-gardien, balance-balcon*.

It turned out that people were faster in detecting *pa* in *palace* than *pal*, and faster detecting *pal* in *palmier* than *pa*, as the following graph illustrates (the horizontal axis denotes the type of sequence the subjects were looking for, the vertical axis how much time it took them to press the button):

(156)



These data show that the French speakers divided the word into syllables while they were listening to them, and found it more difficult to hear sequences that were smaller or bigger than a syllable, although they did manage to do that as well — they were thus not completely restricted by the perception of syllables, but their first guess seemed to always have been to listen to syllables.

Later experiments showed that the same effect does not hold for all languages. In particular, English speakers did not show the effect of syllable structure at all, and found it just as easy to find *pa* as *pal* in any word which contained those sequences. The reason may be that syllable boundaries in English are relatively blurry as compared to e.g. German and French: it is difficult to find agreement about where exactly the boundary is in words like *bitter*. It might therefore be less helpful to people listening to English to pay close attention to syllables than it is for speakers of other languages.

Syllabaries

alphabetic systems

The final type of evidence for syllables we want to discuss are writing systems. You are probably familiar with *alphabetic systems*, in which every letter corresponds (roughly) to a segment — the Roman alphabet is such a system, and this book is written in it, although the correspondence between sound and letter in the case of English is very rough — and *logographic systems*, in which every symbol stands for a meaning unit, such as a morpheme: Chinese is often cited as an example of the latter (although matters are more complicated, especially in Modern Chinese).

See Section ??

logographic systems

However, another quite widespread system is *syllabic writing*, in which the syllable is the smallest unit of writing. One example is (or was) the Akkadian *cuneiform*, a system that was invented by the Sumerians around 4000 BCE and written on clay tablets. The system had several logograms as well, but at its core it was syllabic. Here are some examples:

syllabic writing

cuneiform

(157) 

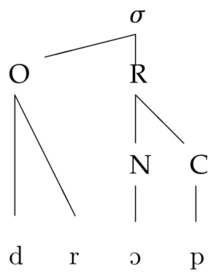
Other examples of languages which have, or used to have, a writing system based on syllables are Ge'ez (Afroasiatic), Mayan (Mayan), Mycenaean Greek (Indo-European) and Cherokee (Iroquoian). Many alphabetic systems seem to have developed out of a syllabic system at some point: when humans invented the art of writing sounds, the first building blocks that apparently came to mind were syllables. In many cases, these building blocks were only very simple syllables, consisting of a consonant followed by a vowel. In the following section, in which we examine the internal structure of the syllable, we will discover why.

5.2 The internal structure of the syllable

Even though we are thus unable to directly observe the phonological syllable, there are plenty of reasons to assume that it has an organizing role in the sound systems of many, if not all, languages of the world. There is also evidence that the structure in (149), of one vocalic head and a few consonants preceding and following the vowel, is too simplistic.

You may have already observed in the preceding section that we sometimes took decisions that require more motivation than we actually gave. For instance, in the French perception experiment we assumed that *pal* is a syllable in *palmier*, but not in *palace*. Similarly, in the Yaqui example in (152), we were tacitly assuming that some consonants belong to the same syllable as the preceding vowel, whereas others do not. This may have seemed quite plausible, because it is the same as one would do in English, but how do we motivate these divisions further?

In this section, we will study the internal structure of the syllable in somewhat more detail. This structure looks as follows (again for the English word *drop*):

(158)


onset
rhyme
nucleus
coda
subconstituents

The capital letters in this graph denote an *onset* (O), a *rhyme* (R), a *nucleus* (N) and a *coda* (C). These *subconstituents* of the syllable are all headed: in *drop*, *d* is the head of the onset, *p* the coda and ɔ the head of the nucleus, the rhyme and the syllable.

Onset

There is one property of headedness in constituents that we have not mentioned yet: heads are obligatory, whereas dependents are not. This means that every syllable in every language always will be expected to have a nucleus and a rhyme, whereas not every syllable necessarily has an onset.

An example of a language without obligatory onsets is Indonesian (Austronesian). Here is how you count from one to ten in this language:

(159) *satu, dua, tiga, empat, lima, enma, tujoh, delapan, sembilan, sepuluh*

hiatus

Most numerals in this language start with a consonant, but the words for *four* and *six* start with a vowel. Since we assume that all words are divided into syllables, this means that at least these words start with a syllable with a rhyme, but no onset. Similarly, the word for 'two' has two syllables, of which the second does not have an onset (a configuration in which two vowels occur next to each other without an intervening consonant is called *hiatus*).

Obviously English is like Indonesian in that it has syllables without an onset — one only has to look at a few words like *eight* and *eleven* to be convinced of that. Many other languages also allow onsetless syllables.

The requirement that syllables have onsets is on the other hand very strong in certain languages. An example of this is Axininca (Maipuran). Whenever the concatenation of morphemes would result in an onsetless syllable, an epenthetic [t] is inserted in this language, as the following examples demonstrate:

(160)	a.	/no-ŋ-koma-i/	[noŋkomati]	'he will paddle'
	b.	/no-ŋ-koma-aa-i/	[noŋkomataati]	'he will paddle again'
	c.	/no-ŋ-koma-ako-i/	[noŋkomatakoti]	'he will paddle for'
	d.	/no-ŋ-koma-ako-aa-i/	[noŋkomatakotaati]	'he will paddle for it again'

epenthesis

(160) shows several affixes which have been added to the base form *koma*. Whenever such a suffix starts with a vowel and the preceding stem or suffix ends with a vowel, a consonant *t* is inserted. This *t* provides the following syllable with an onset. (Such insertion of phonological segments is called *epenthesis*.)

There are thus languages in which onsets seem obligatory, like Axininca, and languages in which they are optional, like Indonesian. There do not seem to be languages in which onsets are consistently disallowed. Since rhymes and nuclei are also allowed (because they are heads), and the minimal form of a rhyme is a vowel (as we will see), we have the following universal:

(161) All languages allow CV syllables.

Syllables with one consonant followed by one vowel are also called *core syllables*. There are no other syllable types which can make this claim to typological universality. Further, core syllables are typically the first that are acquired by children, even in languages that allow other syllable types. Indeed, it has even been demonstrated that at a very early stage of acquisition, children replace even the smaller V syllables with CV structures. The following examples are from Dutch (Indo-European):

core syllables

(162)

adult form	*child's output*	
/oto/	[toto]	'car'
/api/	[tapi]	'monkey'

In these examples, the first syllable of the word has no onset in the adult language. It is filled with a *t* in the child language, just like in Axininca. (The fact that it is a *t* in both cases is a coincidence.)

Certain languages also allow for more *complex onsets*. In many languages, these will consist of two consonants of which the first is an obstruent and the second a sonorant, such as in Sanskrit (Indo-Aryan):

complex onsets

(163) *prath* 'spread', *svaɟ* 'embrace', *mjaks* 'glitter', *mnaː* 'note', *grabh* 'seize'

Sanskrit *v* is probably a glide, and thus a sonorant consonant, as it can alternate with the vowel *u*. The generalisation thus is that onsets of two consonants are allowed, if the second consonant is a sonorant.

The fact that two consonants can occur together at the beginning of a word does not in itself provide sufficient indication that they form a constituent. However, just as there is strong phonological evidence for the syllable , there are also good arguments for taking this position.

See Section 5.1

As a matter of fact, one example briefly discussed above already sheds some light on the issue: the blended form *yout* for *shout* and *yell*. We argued that most blends occur at the boundary between two syllables, but there also are examples where blending happens within a syllable. In many cases, like in *yout*, it then occurs at the boundary between onset and rhyme.

See p. 103

Blending is not always a speech error in English. It can sometimes also be a way in which speakers consciously create new words out of two existing ones:

(164) *smog* (*smoke* and *fog*), *brunch* (*breakfast* and *lunch*), *motel* (*motor* and *hotel*), *infotainment* (*information* and *entertainment*)

Interestingly, also in these cases, the blending occurs exactly at the boundary between two constituents. For instance, in *brunch*, it occurs exactly between the onset and the rhyme; there are no blends of the shape *breanch*, in which only the vowel would be taken from *breakfast* and dropped into *lunch*.

As to the headedness of onsets, the most popular type of evidence comes from reduction processes: if, for some reason, only one of the two consonants is pronounced, it is typically the first one, the head. (Recall that headedness involves the claim that the head is necessary but the dependent is less so.) One piece of evidence comes from language acquisition. In a study on the

cochlear implant

acquisition of children with a *cochlear implant*, i.e. a permanent hearing aid that is implanted close to the ear, it was shown that these children tend to reduce clusters at some stage of their acquisition:

(165)

adult word	*child form*
'clock'	[k^{h}akki]
'frog'	[fɔg̊]
'brush'	[bʊtʃ]

There was a handful of cases in which it was the second consonant that survived [l^{h}ak] for 'clock', but the cases where it was the first segment were an overwhelming majority. Very similar results were obtained with normally hearing children in many experiments as well.

Furthermore, adult languages also sometimes show signs of the same phenomenon. An example is reduplication in Sanskrit. Recall that the morpho-

See Section 5.1 on page 101

logical process of reduplication can copy a syllable. In Sanskrit, the copied syllable is simplified: one of the simplifications is in the onset, which consists of only one consonant in the copy. (Other simplications involve the nature of the vowel, which is of no concern to us now.) This copied consonant is always the head. In the examples, the reduplication adds perfective meaning to the stem; I added a dash between reduplicant and base for clarification:

(166)

base	*reduplicant*
prath 'spread'	*pa-pratha*
svaɟ 'embrace'	*sa-svaɟa*
mjaks 'glitter'	*mi-mjakṣa*
mna: 'note'	*ma-mnur*

The fact that in processes such as this, it is uniformly the first consonant that is preserved, can be described nicely by claiming that it is a head — since that is the only obligatory segment of the onset by definition.

Rhyme

As in the case of the onset, the evidence for the rhyme comes from several different sources; we will discuss only a few. One we have actually already seen

See p. (151)

above, in our discussion of Arabic stress : languages distinguish between light and heavy syllables in the assignment of stress. This distinction can (almost)

branching rhyme

always be described as a difference between a *branching rhyme* (with more than one segment, which is heavy and attracts stress) and a non-branching rhyme (with one segment, which is light). In the case of Arabic, *tab* in *katabti* 'you wrote' has stress because it is heavy, but *ra* in *darabu* is light and therefore unstressed.

Another piece of evidence comes from a phenomenon called *compensatory lengthening*. An instance of this can be found in the history of English. Compare the following Old English words with their German cognates:

compensatory lengthening

(167)

Old English	*German*
go:s	*Gans*
o:þer	*Ander*
so:fte	*sanft*
fi:f	*fünf*
u:s	*uns*

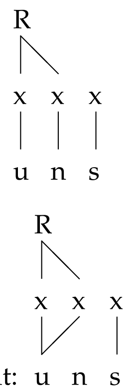
The Old English words all have a long vowel, where the German forms have a short vowel followed by a nasal. There is reason to assume that the latter language is more faithful to the state of affairs in Proto-Germanic, predating all of the Germanic languages, and that English is the language that has changed.

How can we describe what has happened? The assumption of the rhyme gives us a nice tool to formulate this description: the nasal was lost, i.e. it was delinked from its position on the skeletal tier in cases where it was followed by a fricative. After this, the empty position in the rhyme was filled by the preceding vowel. This lengthening was compensatory in the sense that the vowel length compensated for the lost consonant (we disregard the position of the *s* for now 5.4):

See Section 5.4

(168) a. Input:

b. Output:



In this example, the skeletal tier (which we have disregarded so far) has been added explicitly in order to see more clearly what is going on. The crucial step in the argument is that compensatory lengthening only involves consonants that get lost in the rhyme. There are basically no known cases in which compensatory lengthening involves an onset consonant that gets lost and results in lengthening of the vowel. CL is thus restricted to the domain of the rhyme.

The change we are witnessing in (106) is *a priori* different from the phonological processes we have seen before. We are dealing here with a diachronic change, the 'input' in (106) represents some stage in the history of English and the 'output' some other stage; and there is as yet no specific reason to assume that any speaker ever had both of them in his head. Still, diachronic changes like this, too, may give us some insight in the mental representations of speakers. After all, there must be a reason why it is only deletion of consonants in 'heavy' syllables that results in a long vowel.

See Chapter 4

On this assumption, then, we see autosegmental phonology at work. If a long *u*: would be nothing but a sequence of two short *u*'s, we would not really understand what was going on: we would have to say that the nasal would have turned into a full copy of the preceding vowel, which would make the representation of this change rather complicated.

Compensatory lengthening is found in many of the world's languages. A well-known case can be found in Turkish (Turkic). In this language, there is good reason to assume that the process is synchronic, because, depending on sociolinguistic and pragmatic factors, speakers display optional deletion of a consonant (to be more precise, any of /v, j, h/. When deletion occurs, compensatory lengthening follows suit automatically:

(169) a. *kahya* 'steward' [kahja]-[kaːja]
b. *eylül* 'September' [ejlyl]-[eːlyl]
c. *sevmek* 'love' [sevmek]-[seːmek]

Alternatively, one could assume that words are just stored in two different ways in the Turkish lexicon and that one is historically derived from the other. However, such an analysis would disregard the fact that the relation between not having a consonant and having a long vowel is systematic also in the current language (at least for the speakers who have deletion).

Nucleus

Within the rhyme, we sometimes distinguish a nucleus and a coda. The latter (like the onset) is the exclusive domain of consonants, whereas the former is assumed to contain vocalic material, for instance *diphthongs*, sequences of two vowels that occur in the same syllable (with a technical term: they are *tautosyllabic*).

diphthongs

tautosyllabic

A well-known example of a complex nucleus comes from French (Indo-European). This language has words such as the following:

(170) *trois* 'three' [trwa], *croix* 'cross' [krwa], *pluie* 'rain' [plẅi], *truite* 'trout' [trẅit]

These words all start with three consonants: a sequence that we can recognize as a regular onset also in English, followed by a glide [w]. Since there is no place for this [w] in the onset, it has to be postulated somewhere else, and the nucleus seems an obvious place to do so.

Interestingly, there is independent evidence that indeed [w] can function in a nucleus in French. In order to see this, we have to look briefly at the *definite determiner* (the word meaning 'the'). This function word is sensitive to whether the following noun or adjective starts with an onset, as the following examples demonstrate:

definite determiner

(171)

camerade 'friend' [kamrad]	*ami* 'friend' [ami]
le camerade 'the friend' [lə kamrad]	*l'ami* 'the friend' [lami]
les camerades 'the friends' [le kamrad]	*les amis* 'the friends' [lez ami]

If the following word starts with a consonant, the singular determiner has a schwa and the plural determiner has [e]. If the following word starts with a vowel, on the other hand, the singular determiner consists of a consonant [l] only, whereas the plural determiner has a [z] after the [e]. (Notice that the allomorphy is reflected in French orthography for the singular, but not for the plural where *s* is always written even if it does not correspond to anything in the pronunciation.)

We can already understand what triggers this allomorphy: French wants to create an onset in the first syllable of *ami*. In the singular we do this by dropping the schwa at the end of [lə], whereas in the plural we do it by pronouncing an extra [z] after [le].

All of this implies that the definite determiner gives us a test to see whether a noun starts with an onset or not. We can now apply this test to words starting with our glide [w]:

(172)	*watt* 'watt' [wat]	*oiseau* 'bird' [wazo]
	le watt 'the watt' [lə wat]	*l'oiseau* 'the bird' [lwazo]
	les watts 'the watts' [le wat]	*les oiseaux* 'the birds' [lez wazo]

It turns out that there are two types of words in French. Some words, such as *watt*, behave as if they start with an onset, and we should therefore conclude that the [w] is in this onset. Other words, such as *oiseau* behave like *ami* and thus seem to start with a vowel. In these words, then, the [w] can therfore only be part of the nucleus.

It is important that allomorph selection is consistent for the singular and the plural of the definite determiner. There are no words which take e.g. [lə] in the singular and [le] in the plural. This makes it unlikely that speakers just remember what the singular and plural determiners are for every word. Rather, speakers have a system in their minds in which one form of the determiner goes with onset-initial words and another form with non-onset initial words. It is this system that can be neatly described under the assumption of subsyllabic constituency.

Coda

Whereas every syllable has a nucleus and a rhyme, and onsets are at least allowed, and often preferred, in all languages, the coda position is restricted, and dispreferred in many different languages. As a matter of fact, there is an array of languages which do not have coda consonants at all. Examples of these are as typologically diverse as Fijian (Malayo-Polynesian), Mazateco (Mesoamerican) and Cayuvava (isolate). There are no languages in which closed syllables are compulsory. As a matter of fact, the following implicational universal seems to hold:

(173) If a language has closed syllables, then it also has open syllables.

In other words, all languages have open syllables (syllables without a coda), but only a subset has closed syllables.

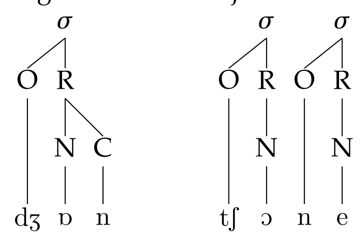
As we mentioned, Boumaa Fijian is an example of a language without closed syllables. In order to repair potential violations of this generalisation, the language employs *vowel epenthesis*, the insertion of a vowel. We can see this at work in loanword adaptation. If a word with a closed syllable is borrowed (from English), a vowel is epenthesized to make the word more well-formed:

vowel epenthesis

(174) Vowel epenthesis in Boumaa Fijian
   a. kaloko ‘clock’
   b. aapolo ‘apple’
   c. tʃone ‘John’

As you can see, various other changes are also performed on these words. For instance, in the first word, the complex onset [kl] is broken up by a vowel, because Fijian does not allow onsets consisting of more than one consonant either. Also, Boumaa Fijian disprefers /dʒ/, and the first consonant of *John* has therefore turned into the voiceless affricate [tʃ]. Another thing you may observe that there are several potential vowels which can be epenthesized. The choice between those will not concern us here, either. The important fact is that final ‘coda’ consonants are not allowed, and vowels are epenthesized to ensure this. In this way, the consonant can be saved by being pronounced as an onset. Here are the syllable structures for the two languages:

(175) English → Fijian



The arrow in this example indicates that the English form is in some way ‘underlying’ to the Fijian form. At some point, speakers of Fijian must have adapted the English word to the phonological system of their language, which does not have codas.

Even in languages that do allow coda’s, the coda position is often quite restricted. For instance, Japanese (Japanese-Ryukyuan) only allows coda consonants if they share their place of articulation with the immediately following consonant. We thus find words such as those in (245a), whereas the forms in (176b) are not allowed.

See Section x

(176) a. *kap.pa* ‘a legendary being’, *kit.te* ‘stamp’, *gak.koo* ‘school’, *tom.bo* ‘dragonfly’, *non.do* ‘tranquil’, *kaŋ.gae* ‘thought’
   b. **kap.ta*, **tog.ba*, **pa.kap*, etc.

This is not exclusive for Japanese; we also find it in an unrelated language such as Ponapean (Micronesian). In this language, we can see that this restriction takes a phonological effect: it causes vowel epenthesis, as the following examples demonstrate:

(177)	/ak-dei/	a.ke.dei	*ak.dei	'a throwing contest'
	/kitik-men/	ki.ti.ki.men	*ki.tik.men	'rat INDEF'
	/naŋkep/	*na.ŋi.kep	naŋ.kep	'inlet'

Another way in which the restriction on codas can be satisfied is by deletion of the offending consonant. Also this is attested in some of the world's languages, e.g. in Jola-Fonyi (isolate):

(178)	/let-ku-jaw/	le.ku.jaw	*let.ku.jaw	'they won't go'
	/jaw-bu-ŋar/	ja.bu.ŋar	*jaw.bu.ŋar	'voyager'

(We leave it as an open question why the final consonant of the word does not need to be deleted; word-final consonants tend to show a slightly different behaviour in many languages of the world.)

5.3 Sonority

The study of the way in which consonants and vowels are arranged in a word is called *phonotactics*. Obviously, dividing the word into syllables, and subdividing syllables in further subconstituents sets one step in the direction of a complete phonotactic theory, with ramifications both for the analysis of individual languages as well as for linguistic typology.

phonotactics

Yet, so far we have been mostly implicit about another necessary step: we need to determine which segments can go into which positions in the onset, nucleus and rhyme: they have some internal ordering. We have briefly and informally mentioned several such restrictions, e.g. that nuclear positions are (typically) vowels, and that languages allow only a small subset of consonants to appear in coda position. Further, we have observed that a typical complex onset consists of an obstruent as the head and a sonorant as the dependent.

In order to make more sense of these observations, phonologists invoke the notion of the *sonority scale*, which looks roughly as follows (there are much more refined versions of the scale, taking into account many more categories, but I think it is very to say that there is at least consensus on this rough version):

sonority scale

(179) *Sonority scale*

obstruents	<	nasals	<	liquids	<	glides	<	vowels
1		2		3		4		5

The notion of sonority was introduced already in 1881, by the German linguist Eduard Sievers (1850-1932), but there is no absolute consensus on the precise phonological or phonetic definition of sonority; as a matter of fact, many different definitions have been proposed. For instance, it has been equated to the openness of the vocal tract and to amplitude (relative loudness). We will ignore the issue here; intuitively, the notion seems to correspond to 'similarity to a vowel'. Obstruents are the absolute anti-vowels. They have a complete or almost complete constriction, they can be voiceless, etc., so they lack almost everything which vowels have. Nasals are already more vowel-like, for instance because they are inherently voiced, although, like plosives, they also involve complete closure of the oral cavity. The constriction of liquids is less constricted, and glides are obviously the consonants that come closest to the vowels.

If we use the numbers in (179) and transfer them into columns of asterisks (so obstruents get a column of height 1, nasals of height 2, etc.), we can represent the syllable structure of the English word *trim* as follows:

(180)

		*	
		*	
	*	*	
	*	*	*
*	*	*	*
t	r	ɪ	m
1	3	5	2

This mountain-like structure is typical of human language syllables. In particular, the vowel is always the highest element (not surprising, given our informal definition; the nucleus is therefore also sometimes called the *peak*). Furthermore, the segments before the peak (thus, those in the onset) gradually rise in sonority, whereas those following the peak, fall in sonority.

peak

There are several points where languages may differ. One is in the required steepness of the rise before and after the peak. Generally, languages prefer to have a rather steep rise in the onset. For instance, in English, [kr] (*creek*) and [kl] are fine clusters, rising from the 1 of [k] to the 3 of [l], but [kn] is not, and neither is any other cluster of an obstruent and a following nasal (disregarding *sn* and *sn* clusters, to which we will return). The reason for this is that the *dispersion* — the difference in sonority — between an obstruent and a nasal is not large enough for the phonotactic grammar of English. Languages can differ on this point: in German, the cognate word *Knie* 'knee' is still pronounced with the initial [kn] cluster, which was lost in English. We can thus say that the minimal dispersion in the English onset is 3 − 1 = 2, whereas in German it can be 3 − 2 = 1. German does allow onsets with greater steepness such as [kr] and [kl] as well, and more generally languages have a minimum restriction on steepness, but no maximum.

See Section 5.4

dispersion

The fall after the peak, on the other hand, tends to be less steep. In many languages, obstruents are not allowed in the coda although nasals and liquids are. The consequence of this is that syllables cannot always be reversed: *calm* is a possible syllable in English, but *mlak* is not, since the difference between *l* and *m* is big enough for codas but not for onsets, We will return to this below
.

See p. 117

The irrelevance of word-edges

We will now take a further look at the structure of the onset. The following words are from Attic Greek, an ancient dialect of Greek (Indo-European) in which the classical playwrights Aeschylus, Sophocles, Euripides and Aristophanes wrote their major works in the 5th Century BCE. Consider the forms in (181), where I have denoted the syllable boundary in each case:

(181)	mi.kron	'small'	ok.to:	'eight'
	pa.tri	'father'	des.mos	'fitting'
	(oi)a.gros	*name*	hag.nos	'holy'

How can we know where these syllable boundaries are correct? Ancient scribes would not even write spaces between words, let alone that they would explicitly mark the boundaries between syllables. Still we have good reason to assume that the syllable boundaries are indeed in this position.

This reason is that the playwrights would write their works in verse, which consisted of regular patterns of light and heavy syllables (in a typical pattern, every line would consist of a number of *dactyls*: either one heavy syllable followed by two light, or two heavy syllables in a row, the latter traditionally called a *spondee*). The first syllables of the words on the left systematically appear in the position of light syllables, whereas those on the right systematically appear in a heavy position. From this we can conclude that the first consonants in the clusters on the left belonged to the onset whereas those in the clusters on the right belonged to the coda. **dactyls**

You will observe that this fits nicely with the theory of the sonority profile lined out above. The complex onset clusters start with an obstruent and are followed by an *r*. On the other hand, the clusters in the column on the right consist either of two plosives or of a plosive followed by a nasal. If Attic Greek had the same dispersion profile as English, then the reason for this is easy to see: a consonant cluster forms a complex onset only if it has the right sonority profile; if not the consonants form a coda-onset sequence (i.e., they are *heterosyllabic*). **heterosyllabic**

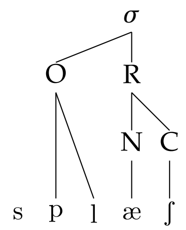
Although this argumentation seems solid, it runs into trouble once we consider the following words:

(182) p^ht^h*ero* 'I destroy', *ksenos* 'stranger', *skapto*: 'I dig'

Exactly the clusters that were avoided word-internally as complex onsets seem to appear here at the beginning of a word. Since there is nowhere else for these consonants to go, the conclusion seems unavoidable that Ancient Greek did have remarkable onset clusters after all.

You may have noticed however, that also English has a number of clusters which do not fit the template that we have described so far. These are words such as *skate, spy, steam, spray, splash, stream* and *scream*. Not only do these words all start with two obstruents, in spite of the high demands on dispersion which English otherwise displays, but words like *spray* and *splash* even start with no fewer than three consonants.

In particular the latter facts give us an indication of what is going on here. All these clusters look like they have an *s* followed by what is otherwise a normal complex onset: an obstruent followed by a liquid. It looks as if in English, a word can be preceded by an *s* that does not belong to the syllable structure proper, so that we have the following structure for a word like *splash*:

(183) 

The reason why *s* possesses this mysterious property in English is unclear (one phonological article about the issue is called "Do you believe in magic?"), but this property of being *extrasyllabic*, i.e. not belonging to the syllable structure might be a property of a number of consonants in Greek. Apparently, the following is true for some languages:

extrasyllabic

(184) At the beginning of the word, at most one consonant (*s*) may be extrasyllabic.

In English and some other languages, extrasyllabicity is restricted to /s/; in Greek, other consonants can have it as well. One might object to this move that it is a trick to save the original hypothesis that onsets consist of maximally two consonants. This criticism is justified to the extent that we do not really understand what is going on here (it is 'magic'), but notice that we have at least restricted the 'trick' to one position in the word.

Furthermore, there is evidence that the 'extrasyllabic' *s* indeed behaves the same as the Greek consonants. This evidence comes from Italian (Indo-European), a language that has onsets very similar to the English ones: normally an onset consists of at most two consonants of which the first is an obstruent, but *str, sp*, etc. are also allowed, in other words, /s/ can be extrasyllabic. In this language, the one syllable in the word which has stress, needs to be heavy . This can be seen in the following facts:

See Chapter 7

(185)
a. *fato* 'fate' [ˈfaːto]
b. *capra* 'goat' [ˈkaːpra]
c. *parco* 'park' [ˈparko]
d. *pasta* 'pasta' [ˈpasta]

Examples (185a) and (185b) show that the vowel is usually lengthened to satisfy this condition, regardless whether one or two consonants follow, as long as the consonants form a well-formed onset. In (185c), you see that this does not happen before a cluster [rk], which has the wrong sonority profile for being an onset. Finally, (185d) shows that /st/ behaves as the non-onset [rk] rather than the onset [pr], in spite of the fact that it can occur at the beginning of the word (as in *stella* 'star').

Notice that this is exactly the same pattern as we saw for Attic Greek before: clusters which are possible at the beginning of the word do not behave as onsets in the middle of the word, but one of the consonants becomes a coda, making the previous syllable heavy.

Interestingly, there is independent evidence for a special position of the /s/ at the beginning of a word. Italian has two forms of the masculine definite determiner (*lo* and *il*), which are distributed in a way that looks very similar to what we saw in French : See p. 110

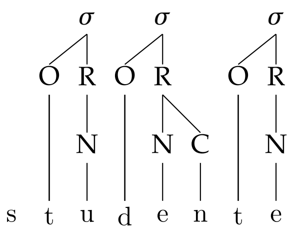
(186) *l'est* 'the East' [lest] (/lo est/) *il burro* 'the butter' [il burːo]
il clima 'the climate' [il kliːma]

We find *il* if the word starts with a (simple or complex) onset, but *lo* if it does not. This gives us the perfect test for *s* clusters' as (187a) demonstrates, these show the expected behaviour, viz. they do not start with an onset, while single *s* does.

(187) a. *lo studente* 'the student' [lo studente]
b. cf. *il senatore* 'the senator' [il senatore]

Clearly, the words do not behave exactly the same as those starting with a vowel (in the latter case the form of the determiner is *l* not *lo*), but the similarity is striking enough to count as evidence that our analysis of the special status of /s/ is right.

The syllable structure of *studente* thus can be drawn as follows, with the *s* outside of syllable structure:

(188)


Syllable contact

Sonority does not only play a role within the syllable, but also across syllable boundaries. In particular, many languages require codas to have a lower sonority than the following onset consonant. For instance, in French, although both [pat] (*pâte* 'pastry') and [ri] (*que je rie* 'that I laugh (SUBJUNCTIVE)' are well-formed syllables, the combination *[pat.ri] is not well-formed; we find [pa.tri] (*patrie* 'fatherland') instead.

We call this requirement the Syllable Contact Law:

(189) *Syllable Contact Law* (SCL) **Syllable Contact Law**
If C_1 is in the coda, and C_2 is the head of the onset of the following syllable, the sonority of C_1 shouldnot be smaller than that of C_2

We already implicitly used the SCL above, in our discussions of Attic Greek and Italian. The SCL can also be seen at work in Korean (isolate). In this language, whenever a sequence arises that would violate the SCL, a phonological process applies to repair this unwanted configuration:

(190)

	Input	*Output*	*Gloss*	*Related form*
Nasalisation:	/sip-nyən/	[sim.nyən]	'ten years'	[sip-ɨl] 'ten-ACC'
	/kam-li/	[kam.ni]	'supervision'	[to-li] 'ethics'
Lateralization	/non-li/	[nol.li]	'logic'	[non-mun] 'research paper'
	/tikɨt-liɨl/	[ti.kɨl.li.ɨl]	'*t* and *l*'	[niɨn-tikɨt] '*n* and *t*

Nasalisation (of plosives and laterals) and lateralisation (of plosives and nasals) apply either to the first or to the second segment in the sequence, depending on certain complicated factors which will not concern us here. What is important for us is that these processes always apply to clusters that violate the SCL and results in clusters that no longer do: the SCL is the *trigger* of the process.

trigger

Neither of these processes occurs if the underlying cluster satisfies the SCL:

(191) a. /kun-tæ/ - [kun-tæ] 'army'
b. /kal-ku/ - [kal.ku] 'desire'
c. /kal-maŋ/ - [kal-maŋ] 'desire'

This is a strong indication that what is at work in these cases is the SCL. Sonority restrictions of this type show up in many languages. Importantly, the reverse effects are never found; these would be languages which have e.g. *rt* onsets, but no *tr*; or which allow *ak.la* syllable contacts, but not *al.ka*.

5.4 The syllable structure of English

To conclude this chapter, we briefly consider the syllable structure of one language. English is a good choice, given that you will be familiar with it when you can read this book, but also because it has a reasonably complex syllable structure. (There are languages which only have core syllables, hence are much simpler, but also languages like Polish, Georgian and Berber, which are much more complex). We will see that the theory developed in the preceding sections gives a good frame for understanding the phonotactics of English, although several details require further elaboration. I concentrate on the consonantal positions here (i.e. onset and rhyme); English vowels are a rather complicated area.

Let us first consider the onset. The constituent is obviously not obligatory in English, given words like *English* and *onset*, which start with a vowel, yet are perfectly acceptable to the English speaker.

Simple onsets can be filled by any consonant, with one notable exception: the [ŋ] can occur at the end of the word (*sing*), but not at the beginning (**ngis*), and, in most varieties of English, not in the onset of a word-internal syllable either. A word such as *finger* is pronounced as [fɪŋgɚ], not as *[fɪŋɚ]. (There are a few marginal cases such as *dinghy*) It is true that *singer* is pronounced as [sɪŋɚ], but this is only possible because there is a morpheme boundary between the [ŋ] and [ə]. (More on the interaction with morphology in Chapter **??** .)

See Chapter ??

As to complex onsets, we can draw the following table (+ denotes that a combination exists, – that it does not or is very marginal):

(192)

C2 / C1	l	r	w
p	+	+	−
t	−	+	+
k	+	+	+
b	+	+	−
d	−	+	+
g	+	+	+
f	+	+	−
θ	−	+	+
s	+	−	+

The first segment of a cluster is always a 'simple' obstruent. Affricates such as [tʃ, dʒ] do not occur in clusters at all (**Jrohn*, **chleaf*). The second consonant in a cluster is either one of the liquids [l, r] or the glide [w]. See Section x

You can see that not all cells are filled in the table. In particular, the combinations [tl, dl, θl] and [pw, bw, fw] are missing. If you observe these triples closely, you will discover that the first three all involve a coronal obstruent followed by a coronal liquid, and the second triple all a labial obstruent followed by a labial glide. In other words, English onsets satisfy the following criterion (hopefully reminding you of the OCP):

(193) The two segments in the onset cannot have the same place of articulation.

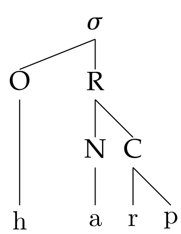
The odd one out is of course coronal *s* which can be followed by a coronal *l*, but as we have already seen , *s* can occupy a position outside of the syllable. The reason why *sr* does not occur in English is a mistery; it seems that in onsets with *r*, English prefers [ʃ] (*shrill*, *shrimp*). See p. 115

We now turn to the coda position in English. We can observe that every consonant can occur in this position, except [h]: *sip, sick, sit, sin, ill, in, rush, rib, kid*, etc. Furthermore, we also find consonant clusters:

(194) *harp, help, lamp, walk, dark, rank, old, word, wound*, etc.

We can take these as evidence for English having a complex coda:

(195)



The sonority profile of the codas in (194) is always the same: a liquid or nasal followed by an obstruent. A coda is therefore almost the mirror image of an onset, except that nasal+obstruent sequences are viable codas (*lamp, rank,* *hand*) but the opposite is not a well-formed onset in English (**pmal*), although it is in other languages.

However, this is not the whole story. In the first place, complex codas are (almost) completely restricted to the word-final syllable in English: a word like **lampgrom* does not seem well-formed: it would not look like an English word if you would first encounter it. In the second place, there are quite a number of apparent counterexamples to the generalization that complex codas have a falling sonority profile:

(196) *act, lapse, past, apt, . . .*

The list becomes even longer if we consider morphologically complex forms (*length, depth, barred*, etc.) Notice that in all these cases, the final, 'offending' consonant is a coronal. It seems that just like the voiceless coronal fricative *s* can be an exception at the beginning of the word, all voiceless coronal plosives can be exceptional at the end of the word. We might say therefore that the set of 'extrasyllabic consonants' at the end of words is larger than that at the beginning of words even though, again, the literature does not provide an answer as to why this is the case.

5.5 Exercises

1. Draw the syllable structure of the following English words: *stream, black, each, blister*.
2. The American poet Adelaide Crapsey (1878-1914) became known for developing a verse form she called the *cinquain*, which was supposed to be an American analogue of the haiku. Below are two examples from her work:

(197) a. **Triad**

These be
Three silent things
The falling snow... the hour
Before the dawn... the mouth of one
Just dead.

b. **Amaze**

I know
Not these my hands
And yet I think there was
A woman like me once had hands
Like these.

Describe the form of the cinquain. Can you give a reason why the existence of such a form (created by a 19th Century American author) is less compelling an argument for the syllable than the South-Slavic decasyllable?

ism

3. A *spoonerism* is an error in which sounds in two different words get transposed. Here are some examples of spoonerisms by English speakers

(198)
a. *three cheers for our queer old dean* 'dear old queen'
b. *is it kisstomary to cuss the bride* 'customary to kiss the bride'
c. *a blushing crow* 'a crushing blow'
d. *a well-boiled icicle* 'well-oiled bicicle'
e. *is the bean dizzy* 'dean busy'
f. *frish gotto* 'fish grotto'
g. *flake bruid* 'brake fluid'
h. *spicky toint* 'sticky point'
i. *The Shaming of the True* 'taming of the shrew' (title of a rock opera)

In what way do such errors provide evidence for subsyllabic constituency? What about speech errors such as *frake bluid*?

4. Italian, like French, has a number of rising diphthongs (e.g. [je, ja, we, wa]. Considering the fact that these rising diphtongs can occur after a single consonant, but never after two in Italian, what is the difference in syllabification between the two languages?

(199)
a. *pieno* 'full' [pjeːno]
b. *chiave* 'key' [kjaːve]
c. *quello* 'that' [kwelːo]
d. *guado* 'ford' [gwaːdo]

5. What explains the epenthesis of ɨ in the following words in Lénakel (Tanna)?

(200)

	Underlying	*Epenthesis*	
a.	/t-n-ak-ol/	[tɨ.na.gɔl]	'you will do it'
		*[tna.gɔl]	
b.	/ark-ark/	[ar.ga.rɨk]	'to growl'
		*[ar.gark]	
c.	/kam-n-m̄an-n/	[kam.nɨ.m̄a.nɨn]	'for her brother'
		*[kam.nm̄ann], *[kamn.m̄ann]	

6. Here are some verbs in present and future tense in Tagalog (Austronesian):

(201)

Present	*Future*
bili 'to buy'	*bibili* 'will buy'
talon 'to jump'	*tatalon* 'will jump'
alis 'to leave'	*aalis* 'will leave'
kain 'to eat'	*kakain* 'will eat'
matulog 'to sleep'	*matutulog* 'will sleep' (you can consider *ma* a prefix)
maligo 'shower'	*maliligo* 'will shower'

How is the plural form derived from the singular? How can this type of morphology be described?

7. In the history of Spanish clusters of consonants sometimes reversed their order. Look at the following examples.

(202)

Latin	*Old Spanish*	*Middle Spanish*	
spatula	espadla	espalda	'blade'
retina	riedna	rienda	'rein'
titulo	tidle	tilde	'tilde'

The relevant clusters have been underlined. What principle of syllabification could explain this change?

8. Voiced obstruents are disallowed in certain positions of the word in German (Indo-European). Consider the following arguments and explain why it can be taken as an argument for the syllable (the examples in the lefthand column are real words of German, the examples on the right could never be).

(203)

a.	[diːp] 'thief'	*[diːb]
b.	[raːt] 'wheel'	*[raːd]
c.	[bɛrk] 'mountain'	*[bɛrg]
d.	[glaːs] 'glas'	*[glaːz]
e.	[motiːv] 'motive'	*[motiːf]
f.	[vɔtka] 'vodka'	*[vɔdka]
g.	[vɪkvam] 'wigwam'	*[vigvam]
h.	[bɛtmɪntən] 'badminton'	*[bɛdmɪntən]

9. It could be argued that open syllables within a word satisfy the Syllable Contact Law in the best possible way. Comment.
10. The oldest known form of Greek (called Mycenaean Greek) was written in a syllabic writing system called Linear B. Some of the properties of this system are that *l, r, m* and *n* are not written at the end of the word or before another consonant: for instance, one wrote *pa-ta* instead of *pan-ta*. Another property was that certain consonant clusters were written as more than one syllable (*po-to-li-ne* for *ptolin*). A third property was that word-initial *s* was omitted before a consonant (*stathmos* became *ta-to-mo*). Comment on each of these three observations from the point of view of syllable theory.
11. English words are nowadays being borrowed into many languages. Sometimes there phonological shape is adapted to the borrowing language. Consider the following words in Brazilian Portuguese (very similar things happen e.g. in Arabic):

(204)

English	→	*Portuguese*
[sl]ide		[isl]ide
[sn]ob		[isn]ob
[st]and		[ist]and

Comment on the reason for this epenthesis from the point of view of syllable theory.

12. In many varieties of English a [j] sound is inserted before an [u] in words like *pew, cue* and *hue*. Give evidence to demonstrate whether the [j] is inserted in the onset or in the nucleus.
13. The online edition of the *WALS* (World Atlas of Language Structures) has, among other things, a chapter on syllable structure, written by Ian Maddison. The map shows three types of languages: with simple syllable structure, with moderately complex syllable structure and with complex syllable structure. Give the templates that are used for each of them, and draw them as in the form of a syllable bracket. (You are allowed to put segments in parentheses.)
14. Look at the following data from Kazakh (Turkic) and explain how they are evidence for the Syllable Contact Law.

(205)	/kol-lar/	[kol.dar]	'hands'	cf. [al.ma.lar] 'apples'
	/murin-ma/	[mu.rin.ba]	'nose-INT	cf. [kol.ma] 'hand-INT'
	/koŋɯz-ma/	[ko.ŋɯz.ba]	'bugs'	cf. [ki.jar.ma] 'cucumber'

15. (If you are a fieldworker.) Working with an informant, give a complete overview of the syllable structure of some language (other than English) along the lines of section 5.4. (Since we have not covered the whole body of phonological knowledge on syllable structure in this chapter, it may well be that you encounter things that do not fit the model of this chapter. Describe those too, and explain what the problem is.)
16. (If you are a computer programmer.) The CELEX database is a database of English (as well as Dutch, German and Tuvan) words, which you can access among other things in syllabified form. Build a program that makes an inventory of all the syllables of English. Do you find any syllables which do not fit the templates described in this chapter?

Sources and further reading

Section 5.1. The text of the Asanaginica can be found on the internet at Wikisources. The data on syllable structure in Quechua (Quechua) are taken from O'Roerke (2008). Watson (2011) gives an overview of stress in (Palestinian) Arabic. The reduplication data from Yaqui are discussed in Haugen (2003). More on the Tagalog language game: Conklin (1956); the Hausa language game is documented in Alidou (1997). A classic study on the role of syllable structure in speech errors is MacKay (1972); his data in the study reported here were from Meringer and Mayer (1895). A classical collection of papers on the relevance of speech errors for phonology is Fromkin (1973). The classical article on syllabification in children is Liberman et al. (1974). The experiment of perception in French was originally reported on in Cutler et al. (1986). More on writing systems and their linguistic analysis can be found in Coulmas (2003).
Section 5.2. The relevance of the Axininca Campa data for phonological theory was pointed out by McCarthy and Prince (1993a). The acquisition data of Dutch children are from Fikkert (1994). The facts about reduplication in Sanskrit are discussed in more detail in Kennedy (2011). The study on complex onsets in children with a cochlear implant is Chin (2006). Old English Compensatory Lengthening has been described in Campbell (1959) and Hogg

(1992) and further discussed in Ewen and van der Hulst (2001); Gussman (2002). The Turkish phenomenon is discussed in Sezer (1986); Goldsmith (1990); Kenstowicz (1994); Gussman (2002). A standard reference to the modern analysis of Compensatory Lengthening is Hayes (1989).

The French examples are discussed in Kaye (1989), who gives them a slightly different interpretation. The Fijian case is analysed in much more detail in Kenstowicz (2007). One study on Japanese (and other) syllable structure is Itô (1986).

Section 5.3 The notion of sonority is discussed in Parker (2002, 2008). It was introduced into phonology by Sievers (1881); the particular scale discussed here is from Clements (1990). The data from Attic Greek are discussed among others in Steriade (1988); Kenstowicz (1994); Kiparsky (2003). The article "Do you believe in magic?" is written by Kaye (1992). The Italian data are discussed in that article as well as in Chierchia (1986) and Davis (1990). Krämer (2009) summarizes the literature on Italian phonotactics. Work on the Syllable Contact Law includes Vennemann (1988); Clements (1990); Gouskova (2003).The Korean data are from Davis and Shin (1999), and the data from Old Spanish are from Holt (2004).

Section 5.4. Some descriptions of English phonotactics are Lass (1976); Harris (1994); Hammond (1999).

General A recent book on syllable structure (presenting a new view) is Duanmu (2009). Hyman (2008) lists and explain all known universals in phonology, including those on syllable structure.

Phonological computation

6.1 Computation and representation

We have so far concentrated on the internal structure of phonological forms: the way phrases and words are organized into smaller units of sound structure. Theories about this aspect are often referred to as theories of phonological *representation*: how are the concrete physical events corresponding to speaking represented in the language system, and in the mind.

representation

Next to this, any phonological theory needs to talk about *computation* as well. Theories of computation talk about how phonological forms are related to each other. For instance, it is not unreasonable to assume that the following words start with something which is the same thing (the same *morpheme*).Here is an example. Turkish is a language which displays the process of *vowel harmony*,[1] which means, roughly, that suffixes take a different form based on the phonological shape of the stem to which they are attached:

computation

morpheme

(206) a. *ip* 'rope', *ip-ler* 'ropes', **ip-lar*, **up-lar*
b. *pul* 'stamp', *pul-lar* 'stamps', **pul-ler*, **pil-ler*

We want to say that the thing which expresses the plural is always the same morpheme (LAR), which sometimes shows up as *lar* and sometimes as *ler*. The relation between *ler* and *lar* is the topic of computational phonological theories. Such theories usually assume that one of the two — say, *lar* — is *underlying* and that the other *derived* from it (for instance by *spreading* a feature from the stem to the base. We tacitly already used such derivational terminology already below.

underlying

derived

Seen in this way — which seems to be the standard view and is the view we will also adopt here — phonological computation thus takes an input, an underlying form, as an input and some more concrete form, a surface representation, as the output. Phonology thus works as a little computer which transforms things which are there in our mental lexicon, where the underlying representations are thought to reside, to things which are closer to the phonetic reality. The latter are derived from the former; since this is the most common view, theories of computation are often also called theories of derivation.

Theories of representation and theories of computation are largely independent of each other, since they are designed to explain different kinds of phenomena. In (206) above, it is a fact about phonological representations in

[1]Kornfilt (1977); Clements and Sezer (1982); Hulst and Weijer (1995); Bakovic (2000).

Turkish that all vowels in a word are either front vowels or back vowels. It is a fact about derivations that it will be the suffixes which change in stem-suffix combinations, and not the stems.

The theory of representations will tell us what are the objects that a phonological theory can talk about — for instance, phonological features, phonological segments, autosegmental tiers, syllables, etc., as well as what kinds of relations these will entertain with each other: it gives us a measure for the well-formedness of an individual representation. The theory of computations, on the other hand, tells us how different representations can be related to each other: when one morpheme takes a different shape in different phonological contexts, how these shapes are related to each other; what are the possible changes a phonological representation can undergo.

6.2 Optimality Theory

Optimality Theory (OT)

In this chapter, we will look at one specific theory of phonological computation, *Optimality Theory (OT)*, which for the past few decades has been dominant in this area, albeit in several varieties (we concentrate here on the most classical, standard variety of the theory).

Operations

minimal operations

Like other derivational theories, Optimality Theory involves derivation from an input form for every morpheme to an output form. We can get from one to the other by applying a number of *minimal operations*, each defined in terms of the representations we use; so autosegmental structures, syllable structures, etc.:

- We can *add* or *delete* a feature – and in this way we may turn /hund/ into [hunt] by deleting the feature [voice] at the end
- We can *spread* features such as tone, or vocalic features such as we have seen for Turkish
- We can *shorten* or *lengthen* consonants and vowels — we have seen an instance of this where we derived *cittá [sː]anta* from a form with an underlyingly short consonant in Italian

Klopt dit?

epenthesize

- We can *epenthesize* (insert) consonants and vowels — such as has happened for instance in the Axininca Campa word [noŋkomati] which is derived from /no-ŋ-koma-i/ ('he will paddle').

Klopt dit?

- We may *change syllable structure* as exemplified e.g. in the derivation between /hund/ (where /d/ is in the coda of the first syllable) and [hun.də] (where it occurs in the onset of the second syllable).

Klopt dit?

- We may change the *stress structure* of the word.

We could apply each and every one of these operations on any underlying form. In actual practice, we only take action however if this improves the form in some sense; if the output form becomes better than the input form.

As we have seen, the reason to delete the feature [voice] in /hund/, is that in this way we can satisfy a requirement on syllable coda's. We can write such a requirement (in OT, these are usually called constraints) as follows:

(207) DEVOICE: Consonants at the end of the syllable should not have the feature [voice].

Notice that this formulation of the constraint presupposes certain representational assumptions, for instance that we can distinguish consonants, that there is syllable structure and consonants can occur at the end of syllables, and that we have a feature [voice]. (37) is probably an instance of a more general principle requiring coda consonants to have as few consonantal features as they can get.

An important step in OT thinking is that we only delete [voice] in /hund/ because DEVOICE asks for it. We would not delete [voice] in e.g. the word *denn* 'then' [dɛn], because there is no constraint which requires deletion in this case. To the contrary, there is a general principle of economy or *faithfulness* — as it is called in OT — which states the following:

faithfulness

(208) FAIHFULNESS: The surface representation should be as close to the underlying representation as possible; do nothing.

If DEVOICE is not operative in German, FAIHFULNESS is; therefore *[tɑn] is a bad output form for /dɑn/; it has violated FAIHFULNESS without necessity.

Two functions

So how do we decide which operations can be applied and which cannot? The (phonological) grammar consists of two functions, called *Gen (Generator)* and *Eval (Evaluator)*. Gen takes an input form and blindly applies phonological operations to it in any conceivable combination. In this way it creates a very large number of possible output forms, called *candidates*.

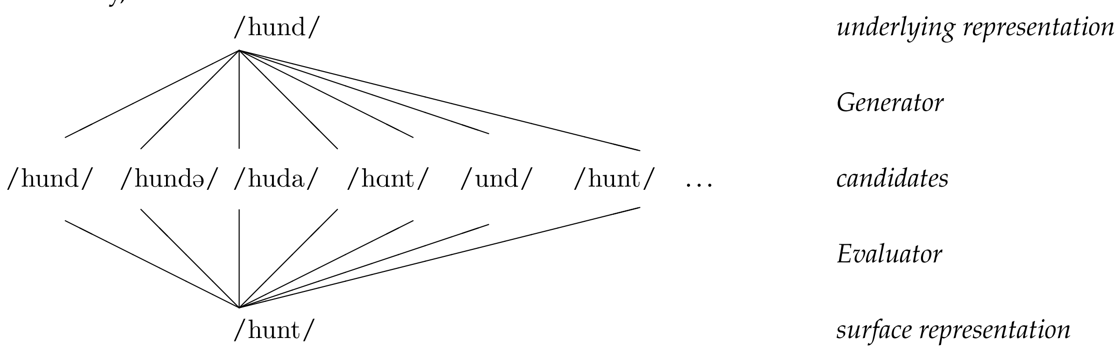
Gen (Generator)
Eval (Evaluator)
candidates

As a matter of fact, this number will be infinite in classical OT. If we take the input form /hund/, we can change all the features of all four segments, but we can also go on adding consonants and vowels to this structure indefinitely. Adding 5,000,000 consonants to this form will probably not improve the structure for any kind of constraint, but Gen is assumed to be blind to this. The idea is not to be psychologically real (nobody assumes that every time you utter a word, you go through all logically possible things you could do to that word), but to offer a precise model of what the best possible form in a language is.

The output of Gen thus is a very large set of candidates. The function Eval takes this set as its input and determines which single one of these best satisfies all grammatical principles, including the two we have just introduced, DEVOICE and FAIHFULNESS. The form is thus a kind of compromise: something has changed, but only in order to satisfy the needs of the language.

Schematically, the derivation can now be drawn as follows:

(209)



In this case, Eval will choose [hunt] as the definitive surface structure for German because it satisfies DEVOICE without making all kinds of unnecessary changes. For the principle of FAITHFULNESS, however, this is not the best possible form; that would have been [hunt].

This observation has a few implications; in the first place, the actual surface form is not perfect in the sense that it satisfies all possible constraints. It is not possible to be perfect in this sense, since constraints can impose conflicting demands. This is why the theory is called Optimality Theory: the 'winner' of the evaluation is not necessarily impeccable, but it is optimal; the best one can do.

Secondly we may observe that apparently DEVOICE has more weight in the grammar of German than FAITHFULNESS. We can write this down as follows:

(210) DEVOICE ≫FAITHFULNESS
(pronounce: 'DEVOICE dominates FAITHFULNESS')

There are also languages in which the order of these constraints is reversed. There is no devoicing in the Yiddish word [hund] and we may assume that the reason for this is that in this language we prefer to be faithful rather than satisfy this particular requirement on syllable structure wellformedness. The crucial difference is that German has the order in (210), whereas Yiddish has the ordering in (211).

(211) FAITHFULNESS ≫DEVOICE

An interesting assumption of (the classical version of) OT is that all constraints are universal; languages differ only in the relative ordering of the constraints. Metaphorically speaking, phonology in all languages consist of a number of forces, and these forces are always the same. The only difference between languages is how powerful each and every one of these forces is. Constraint ranking is the only possible difference between two languages; in this sense, OT is a strong theory of language variation: it claims that systematic differences between languages always can be described with an ordering of universal constraints.

Faithfulness is not one thing

Our analysis of final devoicing in German is not completed yet. It is true that [hunt] satisfies DEVOICE, but this is true also for e.g. [hɑnt] and [hundə]. So why is the former the winner? The answer is relatively easy to give for [hɑnt]. Like [hunt], this form violates FAITHFULNESS, in that it has deleted a feature [voice], but it has done even more: it has also changed the specification for [round] on the vowel, and this is an unnecessary extra violation of FAITHFULNESS. Apparently, we do not just count whether or not a constraint is violated, but also how often this is the case.

Matters are more difficult for the comparison with [hundə]. In order to get there from our underlying form, we arguable need to take only one step: insert an empty vocalic position. So why does this form lose from our winner [hunt]? We will have to split up our cover constraint FAITHFULNESS into a more fine-grained structure of constraints which are all ordered. In particular, we will need at least the following two faithfulness constraints:

(212) a. KEEP-FEATURE: All features in the underlying representation must be present in the surface representation
b. *FEATURE: All features in the surface representation must be present in the underlying representation
c. (Dutch): *FEATURE≫KEEP-FEATURE

We can now consider [hunt] as a better surface structure than [hundə], because the former violates a lower-ranked constraint than the latter.

Tableau

It is common practice to draw the evaluation of surface candidates in a so-called *tableau*; in the case at hand, this tableau looks like this: **tableau**

(213)

/hund/	DEVOICE	*FEATURE	KEEP-FEATURE
hund	*!		
☞hunt			*
hundə		*!	
hɑnt			**!

This should be read in the following way. In the left-hand column you see the underlying representation on top. Immediately below it you see some of the more interesting output candidates. Given that there are infinitely many things we can do, it is impossible to draw all of them, but you do not have to worry about this: it is typically possible to determine which forms are relevant for a discussion and which are not.

From left to right, you see the names of the relevant constraints, in the order in which the grammar (of German, in this case) has ordered them. An asterisk in a cell indicates that the form in question violates a constraint, and two asterisks indicate that it violates the constraint twice. An exclamation mark behind an asterisk indicates that this violation is 'fatal' for the form in question; it is the reason why this form is not the ultimate winner. The pointing finger directs the reader's attention to the form which has no fatal violations and is therefore the optimal form and the actual surface structure.

6.3 A case study: Nasal assimilation

Let us now turn to an example which is slightly more complicated, viz. the behaviour of the English prefix *in-*, which displays nasal assimilation. Nasal assimilation is a phenomenon which is much more wide-spread in languages of the world, and has been analysed in terms of autosegmental representations.

How are we going to integrate such an analysis into an OT framework? We obviously need to have the right constraints, which is a craft in its own right. This section will be an exercise in formulating one such constraint.

Consider the input representation (214a) and the candidate outputs in (214b).

(214) a. /in+polite/
b. { [inpolite], [iŋpolite], [impolite] }

The fact that [impolite] is the winner means that we prefer a structure in which the nasal and the following consonant share their Place node. Let us formulate the relevant constraint as follows:

(215) PLACEHARMONY (first version): A nasal has to bear the same Place node as a neighbouring consonant.

Clearly, *impolite* is the only form among our set of candidates satisfying this constraint, if we assume that it has the following structure:

(216)
ɪ m p o l a j t
Place

However, as soon as we make our set of candidates just a little bigger, we will see that there are more possibilities. Take for instance the form *intolite*, in which the stem consonant has adapted to the prefix, rather than the other way around:

(217)
ɪ n t o l a j t
Place

Presumably, the nasal in the input /in+polite/ is more sensitive to harmony than the plosive is. There are two possible explanations for this difference. It could be that there is some internal difference in the structure of nasals and segments which causes the asymmetry; alternatively, this could point to a difference between affix segments and stem segments. Although the second explanation has something to say for it (as we will see later in chapter), there are also clear indications that there is something right about the first explanation. In English place assimilation, for instance, there is always a nasal involved: there is no assimilation in words such as *actor* (*[ækkɔr], *[ættɔr]). On the other hand, nasals are always homorganic to a following plosive also inside a word (*land, antenna, camp, lampoon, bank, banquette*).

***

Let us assume than for the moment that the second analysis is the correct one. We can build this restriction into our theory in various ways. We could make it a matter of a faithfulness constraint, which would somehow say that nasals tend to be less faithful to their underlying representation than other consonants. Alternatively, we may build the restriction into our definition of the constraint PLACEHARMONY, which we will do here:

(218) PLACEHARMONY (second version): A nasal at the surface structure has to bear the underlying Place node of a neighbouring consonant.

This now explains why /in+polite/ is not rendered as [intolite]. But we are still not completely satisfied. This second version suggests that every nasal will borrow the place of its neighbor, either on its left or on its right. This is not true for English, witness words such as *techno* in which there is no assimilation at all.

Again, there are several possibilities. We could assume, for instance, that the difference with the previous case is that nasal occurs on the righthand side of the neighbouring consonant rather than on its left. We could now revise PLACEHARMONY in the following way:

(219) PLACEHARMONY (third version): A nasal at the surface structure has to bear the underlying Place node of a neighbouring consonant on its righthand side.

This version will work sufficiently well for English, even though the question remains open what is so special about the righthand side of the nasal. Yet if we consider other languages, we discover soon enough that 'righthand side' and 'lefthand side' are not the right concepts to be used. Since OT assumes that constraints are universal, we are however on the quest for a constraint which can explain the facts of as many languages as possible.

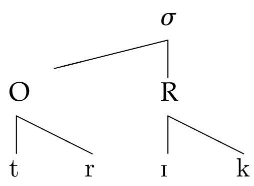
We will have a brief look at Dutch dialects. Many of these dialects — we take the dialect of Hellendoorn, a small town in the north east of the Netherlands, as an example — show syllabic nasals, for instance as the infinitival ending: *eten* 'to eat' [ɛtn̩]. The nasal forms the nucleus of the syllable on its own in cases such as these. Interestingly enough, also this nasal is sensitive to place assimilation, and shows up with the same place as the preceding consonant:

(220) roe[pm̩] 'call'
wer[kŋ̍] 'work'
po[fɱ̍] 'roast'

Syllabic nasals can also borrow their place from their neighbour on their righthand side, for instance if they function as indefinite determiner clitics:

(221) [n̩] doeve 'a pigeon'
[ɱ̍] fietse 'a bicycle'
[m̩] bal 'a ball'
[ŋ̍] keer 'once (a time)'

Apparently, left and right are not the relevant categories, at least not in Hellendoorn Dutch. Still, also in words like *opnemen* 'take on' or *pneumatisch*pneumatic we would not find assimilation in this dialect, showing that not every nasal assimilates in place. The correct definition of the (universal) constraint on Place assimilation is not sensitive to these categories, but instead of this to syllable structure. From your introductory class to phonology you may recall that it is usually assumed that syllables form constituents of the following type (disregarding various details):

(222) 

The generalisation now seems to be that nasals within the coda assimilate, but nasals within the onset do not. An improved version of PLACEHARMONY will therefore say:

(223) PLACEHARMONY (fourth version): A nasal in the coda has to bear the underlying Place node of a neighbouring consonant.

We could speculate why a nasal in the rhyme has this peculiar property. It seems reasonable to relate it to the effect of DEVOICE and other constraints we have seen, which all state that independent consonantal features in the rhyme are undesirable. In some sense, rhymes are the domain of vowels and vocalic material and consonants are aliens in that domain; consonants belong to the onset, where they are much less restricted: they do not have to devoice and they do not have to assimilate. We could now go on to find a general constraint of this sort which will give us all the right results, but that is not something we can go into here.

It is still (intentionally) unspecified in our constraint which of the two neighbours is going to lend its place in case of a choice. Hellendoorn facts shed light on this issue as well:

(224) a. *loop [ŋ] keer* 'walk one time'
b. *(ik heb) de kat* [m] *bettien (gevoerd)* '(i feeded) the cat a little bit'
c. *(ik heb het) rek* [m] *verfien (gegeven)* 'I painted the rack (I gave the rack a little paint)'

In these cases there seems to be a preference for the consonant on the right-hand side. Does this mean we will have to build the notions 'left' and 'right' into our theory after all? An important observation is that in these cases we are considering a determiner which entertains an intimate relationship with the noun on its righthand side and a much less intimate relationship with the word (verb or noun) on its lefthand side. The former is within the same syntactic phrase, but the latter is not. It thus is not necessary to distinguish between left and right; we just have to understand that the nasal attracts place from its closest neighbour in terms of syntactic structure. That could be built into the ultimate version of the constraint as well, but we will refrain from doing that here.

6.4 Beyond place harmony

In our discussion of nasal harmony, the notion of *coda* is important, which we know from chapter 5 (more precisely, section 5.2). We have already seen there that coda consonants are weak, and dispreferred in many different languages.

As a matter of fact, there is quite an array of languages which do not have coda consonants at all. Examples of these are Fijian, Mazateco and Cayuvava, and the following implicational universal seems to hold over known phonological systems:

(225) If a language has closed syllables, then it also has open syllables.

All languages have open syllables (syllables without a coda), but only a subset also has closed syllables.

In section 5.2 we have seen that Fijian is an example of a language without closed syllables. In order to repair potential violations of this generalisation, the Boumaa dialect employs *vowel epenthesis*, the insertion of a vowel. If a word with a closed syllable is borrowed, a vowel is inserted to satisfy the constraint against closed syllables (the following is repeated from example (174) in chapter 5): **vowel epenthesis**

(226) Vowel epenthesis in Boumaa Fijian
a. kaloko 'clock'
b. aapolo 'apple'
c. tʃone 'John'

In order to capture this effect, we can posit a constraint NOCODA:

(227) NOCODA: Syllables should not have a coda, *C $]_\sigma$

Like all constraints, the constraint NOCODA should be assumed to be universal, it is present in all grammars. The difference between English, allowing codas and Fijian, disallowing them, is one in constraint ranking with respect to a faithfulness constraint (ignoring a few segmental differences between the languages):

(228) a. NOEPENTHESIS: (= a subtype of faithfulness) Do not insert vowels
b. English grammar: NOEPENTHESIS≫NOCODA
c. Fijian grammar: NOCODA≫NOEPENTHESIS

(229) a. English

/ʤɔn/	NOEPENTHESIS	NOCODA
☞ʤɔn		*
ʤɔne	*!	

b. Fijian

/ʤɔn/	NOCODA	NOEPENTHESIS
ʤɔn	*!	
☞ʤɔne		*

If we now look at the other side of the syllable template, the onset constituent. The typological behaviour here is quite different. We can posit an implicational universa here as well, but it runs in the opposite direction:

(230) If a language has syllables that lack an onset, then it also has syllables that have an onset.

In other words, all languages have so-called CV syllables, but not all languages have syllables that consist of only a V; we have seen an instance of such a language last week: Axininca Campa. We have expressed this informally before in the observation that the onset is the consonantal domain and the rhyme the domain of the vowels.

In order to describe this situation, we need a constraint of the following type:

(231) ONSET: Every syllable should start with a consonant.

Note that this constraint is almost exactly the mirror image of NOCODA; together they describe the ideal syllable template CV, which all languages have. Formally, the reason for this state of affairs is that no matter how high or how low we rank the faithfulness constraints with respect to these two constraints, CV syllables will always surface, viz. when they are underlying: nothing will have to change to them in order to get to the surface.

Another universal follows from these two constraints:

(232) An underlying (monomorphemic) sequence VCV will be syllabified in all languages as V.CV

(232) is not completely self-evident. It is not hard to imagine a world in which it would not be true. For instance, in French *pat* in *pâte* 'pastry' [pat] and *e* in *é(gale)* 'equal' [e (gal)] are both well-formed syllables, so why do we syllabify *pâté* 'paste' [pate] as [pa.te] rather than *[pat.e]? The answer is that these two constraints conspire to this result:

(233) a.

/pate/	ONSET	NOCODA
☞pa.te		
pat.e	*!	*

b.

/pate/	NOCODA	ONSET
☞pa.te		
pat.e	*!	*

In other words, every language which has *pâté* will syllabify it in the French way. Faithfulness constraints are irrelevant, at least as long as we assume that there is no syllabification in underlying representation (a standard assumption although it is sometimes contested).

Epenthesis

We have seen that both Fijian and Axininca Campa solve their problems with syllable structure by way of vowel epenthesis. We will now go into this a little deeper for the latter language.

The ONSET constraint is very strong in Axininca. Whenever the concatenation of morphemes would result in an onsetless syllable, an epenthetic [t] is inserted:

(234)

a.	/no-ŋ-koma-i/	[noŋkomati]	'he will paddle'
b.	/no-ŋ-koma-aa-i/	[noŋkomataati]	'he will paddle again'
c.	/no-ŋ-koma-ako-i/	[noŋkomatakoti]	'he will paddle for'
d.	/no-ŋ-koma-ako-aa-i/	[noŋkomatakotaati]	'he will paddle for it again'

Once we introduce a specific faithfulness constraint against epenthesis, we have all the constraints set in place to describe this behaviour:

(235) NOEPENTHESIS: Segments in the output should also be present in the input.

(236)

/no-ŋ-koma-i/	ONSET	NOEPENTHESIS
☞noŋ.ko.ma.ti		*
noŋ.ko.ma.i	*!	

However, given the properties of the Gen function, we should also take into account numerous other candidates. Most interesting among these are those forms which satisfy *both* ONSET and NOEPENTHESIS. This is certainly possible; for an input /no-ŋ-koma-i/ there is an output candidate [noŋ.ko.ma] in which nothing is epenthesized, but there is also no ONSET violation.

The point here is that here a different type of faithfulness constraint is violated, viz. one against deletion:

(237) NODELETION: Underlying segments (vowels) must be preserved in the output.

Apparently, this constraint dominates NOEPENTHESIS in Axininca:

(238)

/no-ŋ-koma-i/	ONSET	NODELETION	NOEPENTHESIS
☞noŋ.ko.ma.ti			*
noŋ.ko.ma.i	*!		
noŋ.ko.ma.i		*!	

We have ordered ONSET ≫NODELETION, but it is not very hard to see that we would have got the same result if we would have ordered these constraints in the opposite order, given the fact that NOEPENTHESIS is low ranking (you will be asked to show this in exercise 10). In cases like this, we say that the ordering is irrelevant, which we write down as ONSET, NODELETION, so with a comma instead of the ≫-sign.

More generally, we can order three constraints in 3x2x1 = 6 different logically possible ways: all three constraints can be in the first position, but once we have chosen one, only the two remaining ones can be put in the second position, and if they are fixed, only the one remaining constraint can be put in the final position. Similarly, if we have four constraints, the number of orderings is 4x3x2x1=24, and the number of orderings for five constraints is 5x4x3x2x1=120. These numbers are also written as 3!, 4! and 5!, respectively, in mathematics, which are pronounced 'the factorial of 3, 4, 5' respectively. If we write down all possible orders for a given set of constraints, we get a *factorial typology*. The prediction is that every individual grammar should describe **factorial typology** some (possible) human language.

We thus have six different possible constraint rankings for our three constraints. Yet some of these grammars produce exactly the same result no matter what the input is:

(239) Factorial typology for { ONSET, NODELETION, NOEPENTHESIS }:

1. ONSET, NODELETION≫NOEPENTHESIS: Consonant epenthesis to create onset (e.g. Axininca)
2. ONSET, NOEPENTHESIS≫NODELETION: Vowel deletion to create onset (e.g. Modern Greek)
3. NOEPENTHESIS, NODELETION≫ONSET: Onsetless syllables freely allowed (e.g. English)

We this have three possible different languages, according to this miniature typology. Every language should fit into one of these three categories.

Harmonic bounding

It is one prediction of Optimality Theory that changes never happen without a cause. If we delete something, we violate NODELETION; if we insert something, we violate NOEPENTHESIS. Such violations will only be allowed if they help us satisfy a higher-ranked constraint. Violation of constraints is always *minimal*, because there will always be a competing candidate which has less violations, and unnecessary violation of constraints will not help a candidate in the struggle for life.

minimal

In order to see this, consider the following example from Lenakel. The relevant syllable structure constraint in this language is slightly little different from what we have seen so far, although it is clearly related:

(240) *COMPLEX: Onsets and codas should not contain more than one consonant.

This constraint is responsible for the fact that consonant clusters are broken up by an epenthetic vowel [ɨ] if they would result in syllables with complex marginal clusters:

(241)	a.	/t-n-ak-ol/	[tɨ.na.gɔl]	'you will do it'
			*[tna.gɔl]	
	b.	/ark-ark/	[ar.ga.rɨk]	'to growl'
			*[ar.gark]	
	c.	/kam-n-m̄an-n/	[kam.nɨ.m̄a.nɨn]	'for her brother'
			*[kam.nm̄ann], *[kamn.m̄ann]	

This can be described by assuming the ranking *COMPLEX(, NOINSERTION)≫NOEPENTHESIS for Lenakel. Now study the following alternative candidates for these forms:

(242) a. *[a.rɨ.ga.rɨ.kɨ]
b. *[tɨ.na.gɔ.lɨ]

c. *[ka.mɨ.nɨ.m̄a.nɨ.nɨ]

Like the real winners, all these candidates satisfy *COMPLEX and violate NOEPENTHESIS. The problem is, however, that they violate this constraint more than necessary.

It would be necessary to violate NOEPENTHESIS as often as these forms do it, if Lenakel would have a high-ranking NOCODA, but apparently this is not the case: Lenakel allows closed syllables, so that the language can be assumed to have the following constraint ranking:

(243) *COMPLEX≫NOEPENTHESIS≫NOCODA

Yet even in a language which disallows closed syllables, a candidate such as the following would never win:

(244) *[a.rɨ.ga.rɨ.ɨ.kɨ]

It is safe to assume that this particular form would never win in any language, given the input we studied. It contains an epenthesis which does not improve anything, and it is *harmonically bound* by other forms which do not violate this constraint. **harmonically bound**

Another sense in which vowel epenthesis in Lenakel is minimal is in its choice of the central vowel [ɨ] as the epenthetic vowel. This vowel (as well as its non-high counterpart [ə]) very often serve as the epenthetic vowel. We know why this is: these vowels are quite empty, since they do not contain place features. By inserting them rather than place-bearing vowels, we epenthesize as little as possible into our phonological structure.

6.5 A theory of constraints

You may have noticed that many of the constraints which have been presented here talk about codas in one way or another. Codas are marked positions for consonants. In some languages, they are disallowed altogether, but even in languages which do have them, they are restricted. French word-final floating consonants only show up if there is an onset position created for them, rather than a coda position. Nasals in the rhyme borrow their place features from their neighbour. Obstruents undergo final devoicing in the coda in many languages. We will study a few more examples in this chapter.

Here is the first such an example. Japanese only allows coda consonants if they share a place of articulation with the immediately following consonant. We thus find words such as those in (245a), whereas the forms in (245b) are not allowed.

(245) a. *kap.pa* 'a legendary being', *kit.te* 'stamp', *gak.koo* 'school', *tom.bo* 'dragonfly', *non.do* 'tranquil', *kaŋ.gae* 'thought'
b. **kap.ta*, **tog.ba*, **pa.kap*, etc.

The constraint which is responsible for this is the so-called Coda Condition, well-known from the study of Japanese phonotactics:

(246) CODA-COND: Consonantal place features should occur in a position outside the coda.

Note that the constraint is satisfied by the forms in (245a) under autosegmental assumptions: the place features are all in an onset postion; CODA-COND does not care that they are *also* in a coda. The only structure it militates against is one where place features occur in a coda position to the exclusion of other positions.

The CODA-COND is not idiosyncratic to Japanese; we also find it in an unrelated language such as Ponapean. In this language, we can see that it takes a phonological effect: it causes vowel epenthesis, as the following examples demonstrate:

(247)	/ak-dei/	a.ke.dei	*ak.dei	'a throwing contest'
	/kitik-men/	ki.ti.ki.men	*ki.tik.men	'rat INDEF'
	/naŋkep/	*na.ŋi.kep	naŋ.kep	'inlet'

Another way in which CODA-COND can be satisfied is by deletion of the offending consonant. Also this is attested in some of the world's languages, e.g. in Diola Fogny:

(248)	/let-ku-jaw/	le.ku.jaw	*let.ku.jaw	'they won't go'
	/jaw-bu-ŋar/	ja.bu.ŋar	*jaw.bu.ŋar	'voyager'
	/jaw-bu-ŋar/	*ja.bu.ŋa	ja.bu.ŋar	'voyager'

(We leave it as an open question why it is the first consonant which is deleted rather than the second one.)

We can now see CODA-COND as one member of a 'family' of constraints, all of them having parallel definitions:

(249) a. CODA-COND: Consonantal place features should occur in a position outside the coda.
b. FINALDEVOICING: Consonantal [voice] should occur in a position outside the coda.
c. NASALHARMONY: Nasal place features should occur in a position outside the coda.
d. NOCODA: Consonantal features should occur in a position outside the coda.

In a theory of phonological computation which is based on constraints, such as OT, one should obviously have a theory about what is a possible constraint. If we are allowed to freely formulate new 'universal' constraints all the time, we cannot say that we have much of a theory. We do not make any specific predictions about what is and is not possible in human language, since we can always change the structure of the theory once we encounter a new phenomenon.

Within OT, we posit that all constraints are universal; that is already a restriction of some sort, since we at least need to show how a constraint which we posit for one language plays a role in (all) other languages of the world.

But if we can freely invent constraints, then we can have a constraint X and a different constraint ¬X which says exactly the opposite, and which would 'explain' why we do not see the effect of X in all languages: because many of them would happen to have ¬X≫X, and ¬X would just make X ineffective whenever its occurrence would be unpleasant to us.

Organizing constraints into families such as we have done in (249) is a first step towards building a better theory of constraints. We could build one schematic constraint from which the various concrete instances in (249) can be derived by instantiating the variable F in different ways:

(250) CODA-COND(F): Consonantal feature F should occur in a position outside the coda.

We could now say that the universal set of constraints consist only of concrete instances of a small set of constraint schemes (or even that an individual language chooses one or more instances of the scheme in its actual grammmar.) It is then unexpected that constraints are needed in the analysis of a language which does not fit into some general schema.

6.6 Implications of the theory of computation

Psychological reality

One can see the kind of computation here as reflecting the mapping which speakers probably do while speaking: one retrieves a form from the lexicon, and transforms this in certain ways in order to get to something that can actually be pronounced. The underlying representation stands for the former, the output representation for the latter. Similarly, there must be a mapping going on while listening: the sound waves we here must somehow be mapped onto the structure of words as we remember them.

Phonological computation models some part of this process: it does not deal with actual sound waves, or instructions to the articulatory organs, nor does it deal with configurations of neurons, but it does represent the mapping in some way.

The model of Optimality Theory also abstracts from what is presumably the psychological reality in some other way. We most probably do not entertain an infinite number of possibilities every time we utter a word. In actual practice, the generator function will thus be restricted in some way, and in any case, the theory here seems to be able to describe more *why* a certain input-output mapping is made than *how* it is made.

Altogether, an OT mapping thus gives a fairly abstract account of what is going on; but the claim by its practitioners is that it is also the best, or the most precise account we have. There are various alternatives as well, for instance, people who claim that there is no mapping at all, and that all forms are stored. So German speakers remember both *Hund* [hunt] and *Hunde* [hundə] separately. The fact that one ends in a [t] and the other has a [d], and furthermore that we systematically find no [d] or other voiced obstruent at the end of a German word, is then seen as the result of one or more historical processes, not as something that needs to be represented in the grammar. Under this view, there is no phonological grammar, there is just a collection of words, each of them the result of some path through history.

An argument against this is that people show that they have *knowledge* of patterns such as the devoicing of obstruents in a coda. German speakers display this knowledge in various ways. For instance, it is a mark of a German accent when speaking English to also devoice consonants in that language; but German speakers also do it when borrowing words from English or other languages which do not show final devoicing. Furthermore, in laboratory experiments, Germans will not accept words ending in a voiced obstruents as plausible German words.

All of this implies that even if speakers store both the singular and the plural form for words such as *Hund* 'dog', they also have some way of accessing the regularity of the sound correspondences between these words. From experimental work, we know that speakers do not know about *all* statistically significant patterns which linguists can detect in a language, but only about the ones which somehow make phonological sense, such as the devoicing pattern. When confronted with loanwords, they will only adapt them to such patterns, not to patterns which seem completely random from a phonological point of view. In our terms, they can see patterns which can be expressed by the machinery of phonological computation, but not random other patterns.

This of course implies that we take the cognitive view on phonology seriously. It seems reasonable to say that there is phonological computation, and that it may even be 'optimizing', although the precise way in which it is implemented in the brain may be different from the tableaux we draw in an OT analysis.

Typology

The computational theory of Optimality Theory furthermore provides us with an interesting view on linguistic typology. Remember that the claim is that the *only* differences between languages are in the ranking of constraints, while these constraints themselves, as well the representations about which they are computed are universal. All languages have coda constituents in some sense, but in some languages the constraint against them are disallowed.

Obviously, when we say that the *only* difference in languages is in their constraint ranking, we mean the only systematic differences in their sound system. The fact that the French word for 'tree', *arbre*, sounds very different from the English word, is not the result of the constraint ranking, but from an arbitrary, and hence non-systematic, fact about the French and English lexicons.

Still, the claim that all systematic differences between languages are describable in terms of constraint rankings is a fairly strong one. As we have already pointed out, it means that we make a claim also about universals: every constraint should be present in all languages, even though its effect might be covered by other constraints in some languages. Those constraints should then actually be identifiable, and themselves also universal.

Furthermore, the claim is that every permutation of constraints gives at least a *possible* human language. This hypothesis can then also be tested. For instance, we can check whether we find an attested human language which actually behaves according to the constraint ranking we have established.

If we do find such a language, we find a confirmation of our theory. It is well-known, however, that part of the scientific methodology is to look not for confirmation (or verification), but for falsification of the theory. Unfortunately, it is rather difficult to find such a falsification. The fact that we do not find a language which confirms the expected pattern in itself, could be due for instance to the fact that it is impossible to check all existing languages. But what is worse, the languages we now have in the world almost certainly do not show all *possible* human languages: some of the latter may simply already be extinct without leaving a trace, or yet to arrive in the world, or may even never actualize for extra-linguistic reasons (the people who would speak such a language just give up their language altogether before it could change into the required pattern).

It is important to see, then, that the claim is not about existing or non-existing languages. The claims of cognitively inspired theories of language are ultimately theories about human beings. The claim of Optimality Theory is that humans can compute some languages, but not others. The ultimate test would therefore be to take a given constraint ranking, apply it to some complete lexicon of words, and see whether we can raise a population using this language.

Such an experiment is of course not feasible, but linguistics sometimes use approximations to it, for instance teaching artificial (miniature) languages to (adult) speakers, displaying the required pattern and compare the way in which those speakers acquire patterns that could not be generated by any ranking of the constraints. (More on artificial language learning was said in section 1.2.)

6.7 Exercises

1. Consider the following forms in Yoruba, and provide an analysis in terms of the constraints given in the chapter. Give a constraint ranking and a tableau for the first form, with some of the reasonable candidates. For the purposes of this exercise, you may ignore what happens to tones.

/bu ata/	[bata]	'pour ground pepper'
/gé olú/	[gólú]	'cut mushrooms'
/ta epo/	[tepo]	'sell palm oil'

2. Consider the following forms in Diola Fogny, and provide an analysis in terms of the constraints given in this chapter. Give a constraint ranking and a tableau for the first form, with some of the reasonable candidates.

/let ku jaw/	[lekujaw]	'they won't go'
/ujuk ja/	[ujuja]	'if you see'
/kobkoben/	[kokoben]	'yearn'

3. Consider the following forms in Lebanese Arabic, and provide an analysis in terms of the constraints given in this chapter. Give a constraint ranking and a tableau for the first form, with some of the reasonable candidates.

/ʔism/	[ʔisim]	'name'
/ʔibn/	[ʔibin]	'son'
/ʃigl/	[ʃigil]	'work'

4. Consider the following forms in Samoan, and provide an analysis in terms of the constraints given in this chapter. Give a constraint rank-

ing and a tableau for the first two forms, with some of the reasonable candidates.

[olo] 'rub'	[oloia] 'rub (perfective)'
[sopo] 'go across'	[sopoʔia] 'go across (perfective)'
[aŋa] 'face'	[aŋaia] 'face (perfective)'
[asu] 'smoke'	[asuŋia] 'smoke (perfective)'
[tau] 'repay'	[tauia] 'repay (perfective)'
[tau] 'cost'	[taulia] 'cost (perfective)'

5. Consider the following forms in Turkish, and provide an analysis of the consonant alternation in terms of constraints. You may have to form a new constraint, modeled on the constraints you have seen. Give a constraint ranking and a tableau for the first four forms, with some of the reasonable candidates.

Nominative	**Dative**	
[sap]	[sapa]	'stalk'
[elmas]	[elmasa]	'diamond'
[ev]	[eve]	'house'
[tat]	[tada]	'taste'
[at]	[ata]	'horse'
[deniz]	[denize]	'sea'
[kap]	[kaba]	'container'
[masraf]	[masrafa]	'expense'

6. Draw tableaux for the dataset in (241), taking also the hypothetical forms in (242) into the datasets.
7. Draw a tableau for a hypothetical input /kapta/ in Japanese, assuming it comes out as [kappa] (cf. the dataset in (245)).
8. How many different constraint rankings do we get with 7 constraints? If we call them A, B, C, D, E and F and we assume that (only) the relative ranking of A and B does not matter, in the sense that A≫B always gives the same language as B≫A, how many different languages are produced by these different rankings?
9. There are approximately 7,000 languages in the world. With 8 constraints we can generate more than 40,000 different rankings. Already within this chapter we introduced more than 8 constraints, which however do not suffice to describe all phonological phenomena in languages (let alone all *linguist* phenomena). One might therefore claim that the theory predicts too many different languages. Discuss.
10. On page 135, it is claimed that it does not make a difference for Axininca Campa whether we assume that ONSET≫NOEPENTHESIS≫NODELETION or NOEPENTHESIS≫ONSET≫NODELETION. Show that this is correct, by showing the tableaux for some relevant examples.
11. Korean has a both plain and aspirated stops. Consider the following table, and use an (adapted version of) a constraint from the main text to give an analysis, as well as a typology which includes at least German and English as well.

pat 'field'	*pat*^{h}*e* 'on the field'
tat 'close'	*tat*^{h}*ə* 'to work'
puək 'kitchen'	*puək*^{h}*i* 'in the kitchen'
sak 'old'	*saka* 'to be old'

Are any other languages also predicted by your typology? If yes, what kind of patterns would you find in such languages? (It would be even better if you could name a concrete language, of course.)

12. It has been claimed in this chapter that *VCV* sequences in all languages tend to be syllabified as *V.CV*. However, in English a word like *inadequate* gets syllabified as *in-a-de-quate*. What could the reason be? Invent a constraint which could do the right job. Place it in a constraint hierarchy and draw a tableau to show how your analysis would work.
13. Given an analysis of Belorussian centropetal reduction (in section 2.4 on page 41) in terms of Optimality Theory. You will have to invent your own markedness and faithfulness constraints.
14. It has been claimed in the literature that there are languages which have syllable-final devoicing (as we have seen), but no languages with syllable-initial devoicing. How is this related to our observations on ONSET and NOCODA in this chapter?
15. Suppose we want to test the claim of the previous exercise that there are languages which have syllable-final devoicing (as we have seen), but no languages with syllable-initial devoicing. Set up a small artificial language experiment to study this claim: create two small artificial languages showing the two patterns (but otherwise the same).
(You can now also test these two languages if you have sufficiently large groups of, say, 5 people each, willing to try to learn them. Do they succeed applying the rule on new forms?)

6.8 Sources and further reading

Section 6.1. Anderson (1985b) shows how phonological theory throughout its history has been involved with the study of representations as well as with the study of computation. The subdivision between underlying and surface representations has been introduced by early generative phonology, in particular Chomsky and Halle (1968).
Section 6.2. Prince and Smolensky (1993) is the classic text on Optimality Theory; Kager (1999) gave the first introductory text, nicely summarizing the classical version of the theory up until the moment of the publication of that book for beginning students. Another nice textbook — also talking more generally about doing advanced phonological research — is McCarthy (2008).

Final Devoicing has been the topic of debate for many languages. **?** gives a nice overview.
Section 6.3. Nasal place assimilation is a type of local assimilation, and as such it has been described by **?**. The data from the Hellendoorn dialect are taken from Nijen Twilhaar (1990).
Section 6.4. The Fijian data are from Kenstowicz (2007). The observation that VCV is syllabified as V.CV in all languages is made in various places, for instance in Charette (1991). McCarthy and Prince (1993a) also discuss this, as well as the data on Axininca Campa. The standard source on Lenakel data is Lynch (1974).
Section 6.5. The Japanese data are from Ito (1986), the Ponapean data from Rehg and Sohl (1981), and the Diola Fogny data from Sapir (1965). The idea of constraint families originates with McCarthy and Prince (1993b); I do not think that the particular family organisation here has been proposed in the

literature, but the observation that the various constraints resemble each other has been made more often.

Stress

7.1 Languages with stress

In the previous chapters, we have seen that languages can organize features into segments and segments into syllables. Although it has sometimes been claimed that the arguments for syllables are weaker in some languages than in others, it seems fairly uncontroversial to assume that all languages share this kind of organisation of phonological material.

We could now wonder whether there is also any kind of higher order organisation. Does a speaker of a human language simply utter a string of syllables, one after the other, or are these syllables in turn organized into higher-order units?

At least certain languages seem to give evidence for this higher-order organisation. It seems fair to say that languages tend to organise syllables into words and phrases; these will be the topic of Chapter 8. There are also some languages, such as English, which display an intermediate level of *word stress*: one syllable stands out among the other syllables as being particularly prominent. It is not always precisely determined what the exact phonetic correlates of this 'prominence' are — it is a partly language-specific mixture of higher pitch, longer duration and longer intensity — but speakers will agree which syllable in a word is more prominent. This syllable can be called the *head* of the word.

word stress

head

A proper subset of stress languages also has more organisation of the word, in the sense that some syllables in longer words have *secondary stress*, i.e. they are not as prominent as the head syllable, but more prominent than other syllables in the word. This secondary stress is often rhythmic: stressed and unstressed syllables tend to alternate each other.

secondary stress

Not all languages show evidence for stress, and of those which do, not all languages show also secondary stress. Languages without stress typically have other ways of organizing the word, e.g. by certain autosegmental tonal patterns.

In any case, there seems little doubt that something like a *word* plays a role in the phonological organisation; this word may not always be exactly what people write in between spaces, or as separate characters, but is is remarkably often something coming close to that. In stress languages (on which we will concentrate in this chapter), there furthermore is evidence that such words have a *hierarchical internal structure*: like a syllable has a nucleus, a word has a *head syllable*.

word

hierarchical internal structure

head syllable

Secondary stress furthermore gives evidence for a further level of organisation, in between the syllable and the word: that of the foot. The heads of feet surface as having secondary stress, and the feet themselves have one head, the *head foot*. The syllable which is the head of this head foot has the primary stress. Thus in the English word *encyclopedia*, *pe* is the head of the main foot, and *en* is the head of a secondary foot, whereas all other syllables are unstressed.

A note of warning is in order: it is my experience that stress is sometimes difficult to hear. This is true for people whose native language does not have stress — they are just not used to pay attention to the phonetic cues that stress gives. But even speakers of languages with stress sometimes find it difficult to notice where stress really is. This is not a problem in a textbook like the current one, since I will always indicate stress where this is required, but it is something to be aware of when you are investigating stress in a 'new' language.

7.2 Metrical feet

foot

The notion of a *foot* is derived from classical metrics, the age-old discipline of studying rhythm in verse in classical European languages, in particular Latin and Greek; it has been extended to the study of the rhythmic grouping of syllables within the word. In English poetry, poetic feet are usually bisyllabic, they consist of two syllables. One of these two is more prominent than the other, and this gives us two options:

trochee

In the first option, the first syllable is the most prominent one; we then have a *trochee*:

(251)	s	w	s	w	s	w	s	w
	(Ón	thĕ)	(shóre	stŏod)	(Hí-	ă)-	(wá-	thă)
	s	w	s	w	s	w	s	w
	(Túrned	ănd)	(wáved	hĭs)	(hánd	ăt)	(pár-	tĭng)

(Henry Wadsworth Longfellow, *Hiawatha*)

In this example we see some useful notation illustrated. Accented syllables are denoted by an accent (á), unaccented ones by a breve symbol (ă). Furthermore we put an *s* (for *strong* above the accented syllable and a *w* (for *weak*) above an unaccented one. The brackets indicate that the syllables are grouped in a foot.

iamb

The other possibility is that the second syllable is the most prominent one; we then have an *iamb*:

(252)	w	s	w	s	w	s	w	s	w	s
	Now	is	the	win-	ter	of	our	dis-	con-	tent
	w	s	w	s	w	s	w	s	w	s
	Made	glo-	rious	sum-	mer	by	this	sun	of	York;

(William Shakespeare, *Richard III*)

Iambic and trochaic feet are the most important building blocks in the stress systems of most (stress) languages as well as in poetry. As we have already indicated, feet are different from all other levels of phonological organization (segments, syllables, words) in one important way. Although it is hardly ever contested that all languages have features, segments and syllables, there is quite a number of languages for which there is no evidence for metrical feet; for these languages it cannot be said that some syllable is systematically stronger than its phonological neighbours. We could debate whether this means that these languages have no feet at all, or rather that the feet they have show no phonetic correlates; however, there rarely ever is an empirical difference between thos two claims.

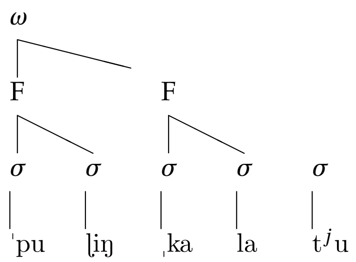
Languages which do have evidence for feet, however, very often choose to have either iambic or trochaic feet. This seems furthermore to be a choice which is made within a language once and for all; there might be no languages in which the two types of feet are mixed.

Pintupi, a Pama-Nyungan language of Australia is a typical example of a language with trochaic feet. In our linguistic transcription, we only note stress, by an accent marker ˈ placed in front of the syllable with primary stress, and ˌ placed in front of a syllable with secondary stress:

(253)	a.	ˈσσ	ˈpaɳa	'earth'
	b.	ˈσσσ	ˈtʲuʈaya	'many'
	c.	ˈσσˌσσ	ˈmaɭaˌwana	'through from behind'
	d.	ˈσσ`σσσ	ˈpuɭiŋˌkalatʲu	'we (sat) on the hill'
	e.	ˈσσˌσσˌσσ	ˈtʲamuˌlɪmpaˌtʲuŋku	'our relation'

The notation which we use in these examples is convenient because it is compact. However, many phonologists really think of these structures in terms of *trees*. The form in (253d), for instance, can be pictured as follows: **trees**

(254)



The straight lines here represent '*heads*' — the most prominent members in a constituent — whereas slanted lines represent '*dependents*' — less prominent members. Furthermore, *F* abreviates 'Foot', and *ω* is often used in the literature for the phonological word. Thus *pu* is the head of the (trochaic) foot *puɭiN*, and this foot is in turn the head of the whole word. For this reason, *pu* gets most stress in the word (it has primary stress), whereas ˈ*ka* (the head of a foot which is not the head of the word, viz. the foot ˌ*kala*) gets less stress (it has secondary stress) and the other syllables get no stress at all. **head** **dependent**

In a word with an odd number of syllables, such as the one in (254), there will be one syllable which does not participate in the foot structure: it is *unfooted*. Languages can choose where this unfooted syllable is located, but usually this will be at one of the two edges of the word: in Pintupi, this is the righthand edge of the word. In other languages, such as MalakMalak (another Australian language, spoken in Western Arnihem), it is the left edge of the word: **unfooted**

(255) a. ˈ*wuru* 'arm, rivulet'
b. *melˈpapu* 'father (emphatic)'
c. ˈ*munanˌkara* 'beautiful'
d. *arˈkiniˌyaŋka* 'we are all going to stand'
e. ˈ*nönöˌrönöˌyunka* 'you (pl) will lie down'

From studying the words with an even number of syllables, it is again easy to see that the language has trochaic feet: stress is always on the first syllable in such words, and then alternates. Furthermore also in MalakMalak, the head foot is the first syllable. However, in words with an odd number of syllables, the languages differ. One syllable is left out of the template, and in MalakMalak this is the first, whereas in Pintupi it is the last. As far as we know, these are the main options; there are no languages which leave for instance the syllable right in the middle of the word unparsed.

We thus have distinguished two axes along which languages may vary:

(256) a. iambic feet vs. trochaic feet
b. first syllable vs. last syllable unfooted in words with an odd number of syllables

This gives us a miniature typology of four different kinds of languages: we expect two types of iambic languages as well as the two types of trochaic languages we have seen. Creek is a famous example in the literature of a language with iambs.

(257) a. *coˈko* 'house'
b. *aˈmifa* 'my dog'
c. *apataˈka* 'pancake'
d. *anokiˈcita* 'to love'
e. *isimahiciˈta* 'one to sight at one'

Again, we can most easily see that we are dealing with an iambic system rather than with a trochaic one in words with an even number of syllables; a word with two syllables simply has an iambic pattern, which is most easily explained if we assume that it consists of one iambic foot. Similarly, also words of four syllables have stress on the final syllable, which makes us assume that there must be two feet there, even though the secondary stress on the first foot is not noted on the data we have.

From the odd-numbered syllable words we can furthermore learn that it is the last syllable of the word which is unfooted. For some reason, this seems to be the option which is chosen by most iambic languages; as a matter of fact, some scholars believe that all iambic languages choose to leave the final syllable unfooted rather than the initial one.

One potential example of an iambic language leaving the first syllable unparsed is Weri, but the data for this language are rather sketchy:

(258) a. *ŋinˈtɪp* 'bee'
b. *kuliˈpu* 'hair of arm'

c. *uˌluaˈmit* 'mist'
d. *akuˌneteˈpal* 'times'

Again, the words with an even number of syllables give us a good indication of the foot structure. They have the same shape as Creek words, although in this case we do have evidence for secondary stress as well. It is the odd-numbered words which show a different pattern, and one which makes us believe that the first syllable might stay unfooted.

It is not clear at present why there are no well attested examples of this language type; given the typology suggested above, we would expect four kinds of languages, but in actual practice only three seem to be attested. Languages seem to prefer to have at least some stress at one of the first two syllables of every word; maybe this is because stress is often used as a *demarcation* of the edges of words. Because the speech signal is uninterrupted, the listener needs cues as to where the word boundaries are in order to be able to make sense of what she hears. Stress can be one such cue that a new word has begun and it has been suggested that this is an important reason why languages have stress to begin with. Allowing words to start with two unstressed syllables might make this task too complex. **demarcation**

Leaving this problem aside for a moment, we now have to translate our different options into constraints on representations. One way of doing this is the following:

(259) a. i. ALIGN(Foot, Left, ˌσ, Left): The left edge of a foot should be aligned with the left edge of the head syllable (so: the heads are on the lefthand side, feet are trochaic).
ii. ALIGN(Foot, Right, ˌσ, Right): The right edge of a foot should be aligned with the right edge of a syllable (so: the heads are on the righthand side, feet are iambic).
b. i. ALIGN(Word, Left, Foot, Left): The left edge of a word should be aligned with the left edge of a foot (so: no unfooted syllables at left edge).
ii. ALIGN(Word, Right, Foot, Right): The right edge of a word should be aligned with the right edge of a foot (so: no unfooted syllables at right edge).

These constraints are instances of some more general family of constraints, aligning phonological and morphological edges to each other. You can see that there would be a general template which these constraints satisfy:

(260) ALIGN(X, Left/Right, Y, Left/Right)

It will be left as an exercise to formulate more instances of this particular constraint schema.

7.3 Syllable quantity

In the languages we have considered so far, all syllables are treated equally. This is a pattern that we find quite often among stress languages. However, in a substantial number of the world's languages, stress is *quantity sensitive* instead: the stress system looks at the structure of syllables and distinguishes between (at least) two types of them: *heavy* and *light* syllables. The distinction is usually connected to the structure of the rhyme in the following way:

quantity sensitive
heavy
light

(261) In heavy syllables, there are (at least) two positions in the syllable rhyme; in light syllables, there is only one position.

Having two positions in the rhyme means having a long vowel, a diphthong or being a closed syllable. Variations on this theme are also possible. For instance in certain languages, syllables are heavy *iff* they are closed by a consonant of a certain type, and light otherwise.

A famous example of a language with a quantity-sensitive system is the Uto-Aztecan language Tübatulabal. In this language, the distinction between light and heavy syllables is made in the following way:

(262) In heavy syllables, vowels are long; in light syllables, vowels are short.

This distinction can be seen as a special case of (261) if we base ourselves on autosegmentalist assumptions on the skeleton.

Consider the following data (we do not distinguish between primary and secondary stress in these examples because we are only interested in foot structure):

(263) a. *iˌponihˌwin* 'of his own skunk'
b. *wiˌtaŋhaˌtal* 'the Tejon Indians'
c. *ˌwɪtaŋˌhataˌlaːbaˌcu* 'away from the Tejon Indians'
d. *ˌyuːˌduːˌyuːˌdat* 'the fruit is mashing'
e. *ˌtaːˌhawiˌlaːp* 'in the summer'
f. *waˌšaːˌgahaˌja* 'it might flame up'
g. *ˌanaˌŋiːˌniniˌmut* 'he is crying wherever he goed (distr.)'
h. *pɨˌtɨpɨˌtɨːdiˌnat* 'he is turning it over repeatedly'

All long vowels are thus stressed. This is the reason why we say that stress is quantity-sensitive in this language: the 'normal assignment' of feet gets interrupted by the requirement that heavy syllables want to be stressed.

In grammatical terms this can be seen as a result of a constraint which is usually called WEIGHTTOSTRESS:

(264) WEIGHTTOSTRESS: Heavy syllables should be stressed.

This constraint has a very high ranking in Tübatulabal grammar — it is never violated. Another observation we can make is that a light syllable before a heavy syllable stays always stressless, whereas light syllables following them are sometimes stressed. This is an indication that we are dealing here with an iambic system: light syllables tend to go into feet with a head on their righthand side. A similar conclusion may be drawn from the first two words, in which there is no heavy syllable at all, and in which the stress pattern is *weak strong weak strong* (abbreviated as *wsws*).

In other words, the language seems to have a basic iambic pattern, and we may assume that also the constraint in (259a-ii), repeated here, is operative:

(265) ALIGN(Foot, Right, ´σ, Right) (henceforth abbreviated as IAMB)

Together, these two constraints will give analyses such as the following, in which we placed feet in between parentheses:

(266) a. (iˌpo)(nihˌwɨn)
b. (wiˌtaŋ)(haˌtal)
c. (haˌniː)la
d. wi(taŋˌha)(taˌlaː)(baˌcu)

The last two examples are not in accordance with the facts of (263): we predict the last syllable in (266c) and the first one in (266d) to be stressless, but this is not the case. They are stressed.

The difference between Tübatulabal and the languages we have seen so far is that in the latter all feet need to be binary: they need to have *both* a head and a dependent. In words with an odd number of syllables, the one 'remaining' syllable stays outside of the foot structure and is unstressed.

For Tübatulabal,a foot can also only have a head and no dependent. Formally, Creek and the other languages have a high-ranking constraint on *foot binarity*: **foot binarity**

(267) FOOTBIN: A foot needs a dependent.

In Tübatulabal, this constraint is dominated by another constraint, which is violated in the other languages:

(268) PARSE-σ: Every syllable needs to be parsed into a foot.

We thus get the following typology for iambic languages (something similar could be done for trochaic languages):

(269) a. Tübatulabal: PARSE-σ≫FOOTBIN
b. Creek (and Weri): FOOTBIN≫PARSE-σ

The difference between the two types of languages wil only be seen in words with an odd number of syllables. In Tübatulabal, the remaining syllable has to be put in a foot, even if that foot is less than perfect as a result. In Creek, one prefers to keep all feet binary, even if that leads to the one syllable being left out of foot structure altogether.

Notice that we can see from examples such as (263d) that FOOTBIN is indeed lowly ranked in Tübatulabal: this word consists exclusively of feet which have only one syllable. The reason for this is of course that all syllables (but the last one) are heavy. In other words, this piece of data provides us with evidence that WEIGHTTOSTRESS≫FOOTBIN, but also that IAMB≫FOOTBIN (because otherwise we could have solved our problem by making the last two syllables of (263d) into one foot). All in all, we thus have established the following miniature constraint ranking for Tübatulabal stress:

(270) PARSE-σ,WEIGHTTOSTRESS,IAMB≫FOOTBIN

7.4 Lexical stress: faithfulness to feet

Word stress in Modern Greek is quite puzzling at first. We may observe that stress can be on many different syllables of the word:

(271) a. *last syllable*: *uraˈnos* 'sky'
b. *penultimate syllable*: *kuˈbaros* 'godfather'
c. *antepenultimate syllable*: *ˈanθropos* 'man'

How are we going to account for this lexical variation? An obvious answer to this is: apparently Greek has feet already present in the underlying representation, and a strong faithfulness requirement on underlying foot structure:

(272) FAITHFOOT: Do not delete underlying feet.

Suppose there are reasons to assume that Greek feet are trochees, and furthermore that *pyjama* (penultimate stress) represents the default. These reasons are manifold; one of them is language acquisition, in which children tend to regularize the other patterns to this one. This gives us the following ranking:

(273) TROCHEE≫FAITHFOOT≫ALIGN(Word, Right, Foot, Right), FOOTBIN

We get the following tableaux for our three example words (leaving out candidates without trochees):

(274) a.

ura(ˈno)s	FAITHFOOT	ALIGN	FOOTBIN
(ˈura)nos	*!	*	
u(ˈranos)	*!		
☞*ura(ˈnos)*			*

b.

kubaros	FAITHFOOT	ALIGN	FOOTBIN
(ˈkuba)ros		*!	
☞*ku(ˈbaros)*			
kuba(ˈros)			*!

c.

(ˈanθro)pos	FAITHFOOT	ALIGN	FOOTBIN
☞*(ˈanθro)pos*		*	
an(ˈθropos)	*!		
anθro(ˈpos)	*!		*

Note that it is not necessary to posit an underlying foot for the default stress structure *kubáros*. This is what it means to be default: the grammar will assign the appropriate structure without instructions from the underlying form. (But note that it would do no harm to assign underlying structure either.)

The constraint TROCHEE is ranked most highly since there is no evidence that there is ever an iambic structure in Greek. Even words such as *uranós* or are analysed as *(ùra)(nós)*.

Still, not everything is possible. One observation to be made is that Greek — like many other languages — displays the effects of a so-called *three-syllable window*: stress is on one of the last three syllables of the word, but never outside it. In other words, (monomorphemic) forms of the following type are unattested in Modern Greek:

(275) *ˈ*makaroni*

The reason for this is straightforward. If we posit an underlying structure *(máka)roni*, the last two syllables are still unfooted. We can then parse these two into a new foot, which will receive primary stress, because this is always on the last foot of the word in Greek. This makes *(máka)roni* different from *(ánθro)pos*, where there is no room to build an extra binary foot.

7.5 The moraic theory of syllable structure

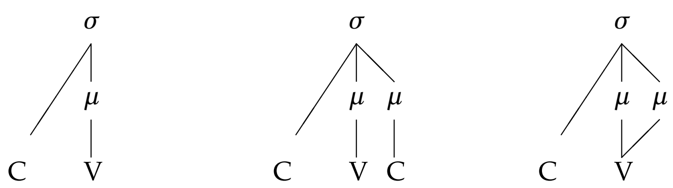
There is a popular alternative to the representations of syllable structure that we have seen so far. Under this conception, the syllable does not consist of an onset and a rhyme, but of two *mora's* (from the Latin word meaning 'a short **mora's**
period of time' or 'delay'). The main generalisation underlying this theory is the following:

(276) a. Heavy syllables consist of two mora's
b. Light syllables consist of one mora

In other words, if both long vowels and coda consonants count, mora's are the same as positions in the rhyme. However, we can also model other kinds of languages using mora's.

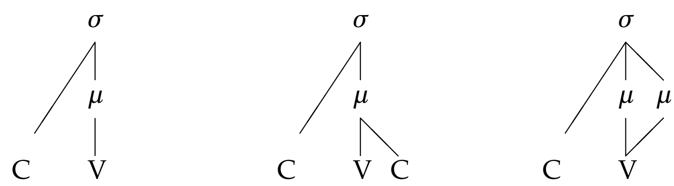
Suppose we are dealing with a language in which closed syllables and syllables with a long vowel are heavy, whereas other syllables are light. We can represent syllable structure in this language in the following way:

(277) a. light b. heavy c. heavy



In a language in which only long vowels count as heavy, on the other hand, we get the following structures:

(278) a. light b. light c. heavy



It is usually assumed that the mora's take the position of skeletal points; the C's and V's in this figure represent root nodes. This means that the phonological timing in this model is slightly different from that in the theory we have developed in earlier chapters, the one based on a skeleton with x-slots: onset consonants do not count for timing, for instance.

One language to which moraic analysis has been applied quite succesfully, is Japanese. As a matter of fact, the mora, called *haku* in Japanese, plays an important role in traditional Japanese linguistics. For instance, in the 'phonetic' part of Japanese spelling, heavy syllables are represented by two symbols, whereas light syllables are represented by one. Traditional Japanese poetry (like *haiku*) is also based on counting 5+7+5=17 mora's (rather than syllables, as in Western renditions of haiku).

It is generally considered true that moraic theory solves two problems of traditional syllable structure/stress analysis. In the first place, weight usually refers to coda consonants and not to onset consonants (as in our discussion of Tübatulabal stress).

In the second place, compensatory lengthening of vowels is claimed to be always the result of the deletion of *coda* consonants, and never of *onset* consonants. So in in our discussion of the history of Germanic in section 5.2 on page 108, we saw that e.g. *gans* can correspond to *ga:s*: the *n* gets deleted and the *a* takes its place. The inverse does not happen: when a consonant in the onset gets deleted in some language, vowel lengthening is not the result (so there are no languages where onset deletion turns *gans* into *a:ns*.

Moraic theory provides us with a formal language which can link these observations and express them in a uniform way. As to weight, we can posit, that some languages build stress feet on syllables, whereas others build them on morae (more on this below). Compensatory lengthening can now be described in terms of mora preservation: if a coda consonant is lost, it may leave a mora behind, which will then be filled by the vowel. However, if an onset consonant is lost, there is no resulting mora, and hence no possibility for onset loss.

It should be noted, that there is some discussion in the literature on the validity of both of these claims. For instance, there are a few languages for which it seems to be true that onsets count for weight. One famous instance of this is Pirahã, which has the following stress rule:

(279) Stress the rightmost heaviest syllable of the last three syllables of the word.

Like many other languages, Pirahã thus displays a three-syllable window at the end of the word: it is as if the stress assignment process only look at those last syllables, Within, this window we choose the heaviest syllable, where the notion of 'heaviness' is defined according to the following hierarchy:

(280) PVV > BVV > VV > PV > BV (> V)
(P = a voiceless plosive, B = a voiced plosive; *a* > *b* means *a* is heavier than *b*)

The notion of weight is thus fairly complex in Pirahã, but it can be decomposed into the following:

(281) a. long vowels are heavier than short vowels
b. syllables with an onset are heavier than syllables without an onset
c. syllables with a voiceless onset are heavier than syllables with a voiced onset

Here are a few examples illustrating these effects (I leave out tone markings and note stress with an acute accent):

(282) a. *ˈkao.ba.bai* 'almost fell'
b. *ˈkaa.gai* 'word'
c. *ˈbii.ao.ii* 'tired'
d. *pia.hao.gi.so.ˈai.pi* ''
e. *ˈʔa.ba.gi* 'toucan'
f. *ʔa.ba.ˈpa* 'Amapá'
g. *ho.ˈao.ii* 'shotgun'
h. *pao.hoa.ˈhai* 'anaconda'
i. *ti.ˈpo.gi* 'species of bird'

(281b) is relevant to our present discussion in particular; it shows that at least in some languages onsets do seem to be relevant to the calculation of syllable weight — there is a handful of languages for which a similar claim has been made.

Compensatory lengthening may also be attested, albeit again in marginal cases. A rather well-known example is the Samothraki dialect of Greek, where deletion of an onset /r/ may result in lengthening of the preceding vowel:

(283) /roɣa/ → [oːɣa], /riz/ → [iːz], /rema/ → [eːma], /roɣa/ → [oːɣa], /ruxa/ → [uːxa], /rafts/ → [aːfts]

(In spite of the phonological notation, it may not be clear that we are dealing with a synchronic process in this case; the underlying representations here as a matter of fact represent Standard Greek and other dialects, but we have no a priori evidence that these are also the underlying representations for Samothraki.)

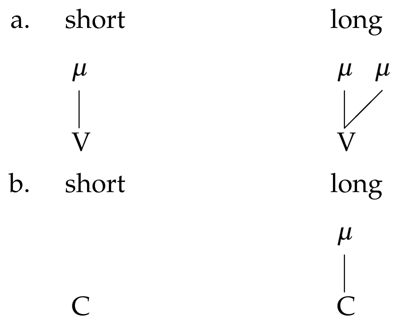
What are we going to do with this type of evidence? A reasonable first approach might be to be very sceptical about it: if our theory forbids it, and the data are so rare, maybe there is something wrong with the sources we have.

However, in this case, this line of attack will not work. In the first place, the Pirahã data stem from afield worker who has spent a large amount of time on his work on this particular language. The Greek dialectological data might be a bit more shaky, but they have been confirmed by some other speaker. In the second place, it is not really true that our theory 'forbids' these facts; there is nothing very deep inherent to any of the theories presented thus far which would disallow onsets to carry morae.

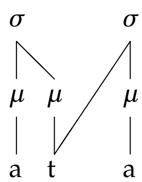
But if this is the case, we are dealing with a typological puzzle: why are data of the Pirahã/Samothraki Greek type so rare as compared to similar effects with coda's? The answer to this might fall outside of the domain of formal linguistics proper: it might have something to do with the phonetic perceptability of codas vs. onsets, for instance.

But given all this, it becomes less clear that mora theory is really superior to the more traditional theories we have seen. If the two reasons why it is introduced in the first place do not really seem to fall within the realm of formal phonological analysis proper, mora theory mainly becomes a convenient notation to talk about the interaction between syllable structure and stress.

As an example of such a notational property, consider the following. The representation of short vs. long vowels will be as in (284a) within mora theory; the representation of short vs. long consonants will be as in (284b):

(284)


One observation which is nicely represented by these pictures is that geminate (long) consonants do not occur in onsets — although, again, there seem to be a few exceptions. Typically, a long consonant will be attached to the coda of one syllable, and the onset of the next one:

(285) [atːa]


monomoraic or bimoraic syllables

Most languages allow for only *monomoraic or bimoraic syllables*; syllables with one or two morae. This means that long vowels could not be followed by long consonants. The following example is from Koya:

(286) /keː tː oːɳɖa/ [ketːoːɳɖa] 'he told'
/oːtːoːɳɖu/ [otːoːɳɖu] 'he brought'

These facts can be understood under assumption of the representations in (285), plus a requirement that Koya syllables have at most two morae, and the idea that nongeminate consonants are never moraic in Koya — this explains why the vowel before the cluster [ɳɖ] does not have to be shortened. Note that especially the latter fact is more difficult to express in a nonmoraic framework.

7.6 Stress typology

Determining where a language fits in our typology

In this chapter, we have introduced a simple model of stress typology. The core of the foot typology are a number of binary distinctions:

(287)
a. Feet are left-headed (trochaic) or right-headed (iambic)
b. Feet are necessarily binary / can be unary
c. Unfooted syllables (if any) or unary feet (if any) appear on the left / on the right
d. Feet are quantity sensitive / quantity insensitive
e. The head foot appears on the left / on the right
f. Stress is lexically determined / predictable
g. Stress is determined on the whole word / on the last three syllables

Because each of these properties is binary, this gives us a rather vast space of possibilites. If we want to classify an individual language, we have to navigate this space somehow. It is important for this to have enough words, of various lengths and with various different types of syllables.

A good way to proceed is as follows. First you try to determine whether the language has iambs or trochees. For this you collect as many words with an even number of syllables as you can find. If they all have a uniform stress pattern, stress is probably *not* lexically determined and also *not* quantity-sensitive. It becomes then fairly easy to see whether the feet are iambic or trochaic.

If the even-numbered syllables word do not show equal stress on all words, you should first check whether stress is always on syllables which are heavy, i.e. whether they have a long vowel or are closed by a consonant. If it looks like this is indeed the case, the best is to first concentrate on even-numbered words with only light syllables. This will tell give you more information about what the 'normal' foot structure is. You can then assume that the feet in words with heavy syllables will be of the same basic type (iambs or trochees).

If it is not clear that heavy syllables attract stress and if the stress seems even randomly distributed in words with an even number of light syllables, the language probably has lexically determined stress. If that is the case, the analysis basically stops, although you may still want to check on long words, whether stress is always at least on one of the last three syllables, or occurs elsewhere.

If stress is not lexically determined, you can then proceed to words with an odd number of syllables. First you check how many stressed syllables there are. Suppose the word has n syllables; if the number of stresses is $(n-1)/2$ (so for instance a word with 5 syllables has 2 stresses), it seems likely that the language only has binary feet, and the remaining syllables are unfooted. Because you already know what the type of feet are, you should now be able to see whether unfooted syllables appear on the left or on the right.

If words with an odd number of syllables n have $(n+1)/2$ stressed syllables (so a word with 5 syllables has 3 stresses), it is more likely that the language allows unary feet. Again, given that you know what 'normal' binary feet are like, you should be able to determine where this 'extra' unary feet is placed.

The only final parameter you have to determine, regardless of what the foot type is, and whether or not the language is quantity-sensitive or lexically determined, is which of the feet gets the main stress. You may safely assume — at least in the case of the exercises to this chapter! — that this is either the first or the last foot. Other options have sometimes been shown in languages of the world, but they are rare and will not be taken into account here.

Some typological gaps

Hayes (1987, 1995) claims that one of the four basic types of feet (iamb vs. trochee, quantity-sensitive vs. quantity-insensitive) which we would expect to exist is typologically inexistent: there are no quantity-insensitive iambs. On the other hand, most trochaic languages seem to be also quantity-sensitive.

Hayes connects this to a psycholinguistic finding (in particular Woodrow, 1909). If we expose informants to a signal *ta-ta-ta-ta-ta-ta-...*, and we alternate the intensity of the *ta*'s, listeners will tend to group them in a trochaic fashion; that is to say, they will tend to hear *...(táta)(táta)(táta)....* On the other hand, if we keep the intensity constant, but alternate the length of the vowels, the listeners will tend to group the sounds as *...(tataː)(tataː)(tataː)....* The conclusion of this is that the difference between foot types is partly determined by the *Iambic/trochaic law* (Bolton, 1894):

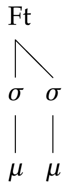
Iambic/trochaic law

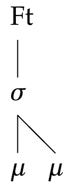
(288) *Iambic/trochaic law*:

a. Elements contrasting in intensity naturally form groupings with initial prominence (trochees).
b. Elements contrasting in duration naturally form groupings with final prominence (iambs).

Trochees now should be constituents which consist of two elements with roughly the same duration. There are two types of these, according to Hayes: we can build feet on the basis of morae, or on the basis of syllables. In the former case, we have a type of quantity sensitive system:

(289) *Moraic trochees*



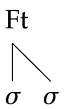


In a system with moraic trochees, heavy syllables will form a foot of their own, whereas light syllables will be grouped together. An example of this is so-called 'Egyptian Radio Arabic', also called 'Cairene Arabic'. In bisyllabic words, stress is on the last syllable if it is (super)heavy, and otherwise it is on the first syllable:

(290) a. *Last syllable (super)heavy: saˈlaːm* 'peace' *diˈmašq* 'Damascus'
b. *Last syllable light: ˈmalik* 'king' *ˈhuna* 'here'

Another possibility is to build trochees on syllables, disregarding the internal structure. We then get a quantity-insensitive trochaic structure:

(291) *Syllabic trochees*

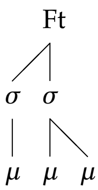


An example of this is Icelandic, where primary stress is on the first syllable of the word, and secondary stress alternates:

(292) *hˈöfðingˌja* 'chieftain (gen.pl)', *ˈakvaˌrella* 'aquarelle', *ˈbíógˌrafíˌa* 'biography'

Yet in iambs, the requirement is that the two parts of the foot are uneven in length, and we have only one canonical foot type:

(293) *Iambs*



An example of this is Tübatulabal, the language we have discussed already. It can furthermore be observed that many languages which use iambic feet have some rule of lengthening vowels and/or consonants to satisfy requirements on foot structure. An example of this is provided by Menomini, a Central Algonquian language. In this language, when a word begins with two light vowels underlyingly, the vowel of the second syllable is lengthened; this can be understood if we assume that these first two syllables are grouped into an iamb:

(294) a. /ahsamaːw/ → [ahsaːmaːw] 'he is fed'
b. /netahsamaːw/ → [netaːhsamaːw] 'I feed him'

It thus seems that we would have to relax our typology of feet some more to also include a preference for uneven trochees.

More in general, it seems that the typology is more lenient for trochees than for iambs. We could also observe that the number of attested (and well-understood) trochaic systems is much larger than the number of attested iambic systems; the latter mainly consist of native languages of (North) America. We could, once again, wonder which conclusions we can draw from these typological considerations. On the one hand, some might wish to argue that the relative paucity of iambic systems is just some accident of history, and that, given this arbitrary historical fact, it is no wonder that there is less diversity in iambic systems: even syllabic or moraic iambs might be possible in principle, but we simply have a much smaller opportunity of finding them actually attested.

Alternatively, some have argued that there is a more principled reason why iambic systems are so few. We could claim, for instance, that iambic feet are not part of our inventory of possible structures. Iambic languages would then need an alternative analysis.

7.7 Exercises

1. Consider the following examples from the Pacific language Awtuw. What kind of foot does this language have — iambs or trochees? Which foot carries main stress?

   i *ˈki.nik* 'sit'
   ii *ˌow.ti.ˈka.yæn* 'old'
   iii *ˈwa.ru.ke* 'big'
   iv *la.pe* 'village'

2. Consider the following examples from the Semitic language Modern Hebrew. What kind of foot does this language have — iambs or trochees? Which foot carries main stress?

   i *ğa.ˈdol* 'big'
   ii *bi.ˈra* 'capital city'
   iii *ta.ˈam* 'tasted'
   iv *me.ˌvu.ga.ˈrim* 'adults'

3. Consider the following examples from the Austronesian language Malay. What kind of foot does this language have — iambs or trochees? Is the language quantity-sensitive or quantity-insensitive?

   i *sən.ˈdar* 'to snore'
   ii *sə.ˈma.di* 'concentration'
   iii *lak.ˈsa.na* 'quality'

4. Consider the following examples from the Malaccan Creole of Portuguese. What can you say about the stress system of this language?

i *ˈka.za* 'house'
ii *sa.ˈba.na* 'fan'
iii *ˌka.za.ˈmin.tu* 'wedding'
iv *ˌmin.ˈti.ra* 'lie'
v *ˌo.rə.ˈsaŋ* 'oration'
vi *ˌkar.ˈbaŋ* 'coal'

5. Consider the following examples from the Palestian Arabic. What can you say about the stress system of this language?

i ˈʃa.[redacted]*a.ra.tun* 'a tree'
ii *ˈka.ta.bu* 'they wrote'
iii *duk.ˈkaːn* 'shop'
iv *baː.ˈbeːn* 'two doors'
v *ˈbaː.rak* 'he blessed'
vi *ˈbaː.ra.ko* 'he blessed him'
vii *ˈka.tab* 'he wrote'
viii *ma.ˈkaː.ti.bi* 'my offices'

6. Give OT tableaus for the derivation of stress in three of the TÜbatulabal examples in (263) (you may choose your own words, except that you may not choose both of the first two, since these have the same pattern).
7. In this chapter it has been claimed that there are no languages which have both iambs and trochees. Show that the OT constraints presented here actually predict otherwise. Discuss.
8. Consider the following examples from the Indo-Iranian language Pashto. What can you say about the stress system of this language?

i *ˈgu.ta* 'knot'
ii *gu.ˈta* 'pochard'
iii *ˈvu.lam.be.də* 'he took a bath'
iv *tʃər.ˈgu.ɾay* 'baby chick'
v *sto.maːn.ti.ˈaː* 'fatigue'

Give a constraint ranking within OT, and draw tableaux for the first two examples.

9. Reformulate the following constraints as special instances of the schema (260) on page 149:

- ONSET
- NOCODA

10. Consider the following examples from Hixkaryana (Carbib); try to place the language inside the typology of moraic vs. syllabic trochees and iambs:

i *ˌow.to.ˌhoː.na* 'to the village'
ii *kha.ˌnaː.ˌnih.no* 'I taught you'
iii *ˌtoh.ˌku.ˌrʲeː.ho.ˌnaː.ha.ˌšaː.ka* 'finally to Tohkurye'

11. Consider the following examples from Fijian (Austronesian); try to place the language inside the typology of moraic vs. syllabic trochees and iambs:

i *a.ˈto.mi* 'atom'
ii *ˌndai.ˌre.ˈki.ta* 'bazaar'
iii *ˌndi.ko.ˈne.si* 'deaconess'
iv *ˌmbeː.ˈle.ti* 'belt'
v *ta.ˌrau.ˈse.se* 'trousers'
vi *mba.ˈsaː* 'bazaar'

12. Consider the following examples from Cahuila (Uto-Aztecan); try to place the language inside the typology of moraic vs. syllabic trochees and iambs:

i *ˈta.ka.ˌli.čem* 'one-eyed ones'
ii *ˈtáx.mu.ˌ* 'song'
iii *ˈqaːn.ˌki.čem* 'palo verde (pl)'

13. Try to give an analysis of Cahuila not in terms of mora's, but in terms of the typology at the beginning of section 7.6. Where does the analysis fail? Can you solve this by using ranked constraints?
14. StressTyp2 is a typological database collecting information about stress patterns in hundreds of languages. The database uses a way to encode stress which is slightly different from the one used in this book, but you should be able to understand it. Try to find example languages for the following patterns:
    - A language with iambs and main stress on the last foot.
    - A language with quantity-sensitive trochees.
    - A language with lexicalized stress within a three-syllable window at the end of the word.

7.8 Sources and further reading

Section 7.2. (Hayes, 1995) is a classical text on metrical stress theory; the Pintupi, MalakMalak, Creek and Weri data are also from that text.

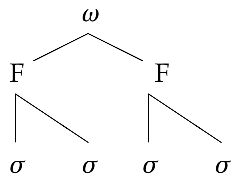
(Topintzi, 2006)

Prosodic structure

8.1 The phonological tree

We have seen in previous chapters that phonological words can be represented by a tree structure: segments are organized into syllables (with some internal structure), syllables into feet and feet into words. We can draw this as follows:

(295)



This tree structure is usually called *prosodic structure* in phonological theory. It is also usually believed that prosodic structure does not end at the level of the phonological word, and that trees reach higher than this. There is variation in the literature as to which higher-order levels are actually present, but it is usually assumed that these involve at least the *phonological phrase* (*ϕ*), corresponding roughly to major syntactic constituents such as (large) NPs or the main predicate VP, the *Intonational Phrase* (IP), corresponding roughly to sentences, and the *Utterance* (Utt), corresponding, well, to a whole utterance of a single speaker.

prosodic structure

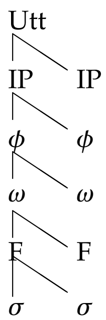
phonological phrase

Intonational Phrase

Utterance

We thus get a tree structure such as the following:

(296)



The hypothesis that linguistic utterances are organized in this way gets support from various different types of evidence, as we will see in this chapter: these structures play a role in our understanding both of phonological and of morphological phenomena which seem to refer to them. For instance, we will see that certain types of phonological processes only work *within* certain prosodic categories: they cannot cross the boundaries of such categories. On the other hand, there are also processes which function only *across* the boundaries of phonological structures.

Notice that the tree structures depicted in (296) mimic those which are used within syntactic theory, but they are simpler in a number of ways: there is a fewer number of different labels — especially given that the only constituents that roughly correspond to syntax are the word, the phonological phrase and the Intonational Phrase —, and secondly, they are not recursive in the same way as syntactic structures. In syntax, we can have for instance, a sentence containing another sentence ('[John admits that [he likes milk]]').

recursion

Such *recursion* according to most scholars is not found in phonological representations: a phonological phrase always directly dominates phonological words, not other phonological phrases.

8.2 The Phonological Word

In the previous chapter, we have already seen one type of evidence for the existence of a phonological word: the existence of primary stress. In some languages, there is one syllable which is clearly more prominent than all other syllables within some domain. That domain does not correspond to the whole string of syllables spoken by the speaker (that would be the Utterance), or a longer stretch of it (that might be for instance the phonological phrase), but to something roughly corresponding to what we might also call a word by other criteria (for instance morphosyntactic ones).

But there is evidence for the existence of such a constituent also from other

/s/ voicing

points of view. A well-known example of this is so-called */s/ voicing*, which we find in nothern Italian dialects. Consider for instance the following cases:

(297)
a. *isola* 'island' /isola/ [izola]
b. *case* 'houses' /kas+e/ [kaze]
c. *amo Sandra* 'I love Sandra' /amo+sandra/ [amosandra]
d. *asoziale* 'asocial' /a+sotsiale/ [asotsiale]
e. *toccasana* 'cure-all' /tok:a+sana/ [tok:asana]

Underlying /s/ changes into [z] when it occurs intervocalically, as you can see in (297a). From data like that in (297b), we know that this also sometimes happens when one of the vowels belongs to a different morpheme — in this case a plural ending.

The domain on which the process is defined is thus bigger than the morpheme. However, it does not apply just anywhere as the following examples demonstrate. In (297c), the first vowel belongs to a different word than the *s*. This means that it is somehow 'too far away' for it to be visible. The /s/ is between two vowels, but it does not voice to a [z].

The domain is thus bigger than a morpheme and smaller than a (syntactic) phrase. This leads us to suspect that it applies within a word. However, the forms in (297d) (a prefixed word) and (297e) (a compound) show that this 'word' does not conform exactly to our understanding of a morphosyntactic word. Both of these are definitely simple words from this perspective, but *s* voicing applies in neither of them.

This is where the notion of a phonological word comes in, a constituent that is similar to that of a morphosyntactic word, but not always exactly congruent with it. In this case, we have to posit that the separate parts of a compound each form a separate 'phonological word', as do prefixes. Otherwise, the boundaries of morphosyntactic words correspond with those of phonological words. The phonological structures for the forms in (297) are thus as follows, where brackets indicate phonological word boundaries:

(298) a. [(izola)]
b. [(kaze)]
c. [(amo)(sandra)]
d. [(a)(sotsiale)]
e. [(tokːa)(sana)]

The phonological analysis can now refer to these phonological constituents. For instance, we may imagine that the constraint responsible for *s* voicing has the following shape:

(299) *(…VSV…): Avoid voiceless coronal fricatives between two vowels within the same phonological word.

It is understood that markedness constraints always are defined on a certain domain, maybe not universally but on a language-specific basis. It is also understood that such domains are always phonological, i.e. that phonological markedness constraints do not refer to morphosyntactic words or phrases.

Of course the way in which a phonological word is constructed itself can be the result of some constraints. These constraints are called *alignment constraints*: they make sure that the edges of morphosyntactic words correspond to the edges of phonological ones. For instance in our case, we have constraints such as this one:

alignment constraints

(300) a. ALIGN(X^0, L, ω, L): The left edge (L) of a morphosyntactic word (X^0) should correspond with the right edge (L) of a phonological word (ω).
b. ALIGN(X^0, R, ω, R): The right edge (R) of a morphosyntactic word (X^0) should correspond with the right edge (R) of a phonological word (ω).
c. ALIGN(ω, L, X^0, L): The left edge of a phonological word should correspond with the left edge of a morphosyntactic word (ω).
d. ALIGN(ω, R, X^0, R): The right edge of a phonological word should correspond with the right edge of a morphosyntactic word (ω).
e. ALIGN(Prefix, R, ω, R): The right edge (R) of a morphosyntactic word (X^0) should correspond with the right edge (R) of a phonological word (ω).
f. ALIGN(Suffix, L, ω, L): The left edge (L) of a morphosyntactic word (X^0) should correspond with the right edge (L) of a phonological word (ω).

g. …

By ranking some of these constraints, we can get the right results in (298). I leave this as an exercise to the reader (exercise 6 on page 165).

In the case of northern Italian dialects, the phonological words are always smaller than morphosyntactic words or they have the same size. It has been argued that we can also have phonological words that are (slightly) bigger than morphosyntactic words. Examples of this we similarly find on the Italian peninsula, but more to the south, for instance in Lucanian dialect.

clitics

The relevant elements in this dialect are so-called *clitics*, small pronoun-like elements that behave morphosyntactically as if they are (more or less) independent words, but that phonologically seem to get integrated with the verbal stem.

The crucial evidence comes from stress. Lucanian has trochaic feet and assigns stress by default on the antepenultimate syllable. The same is actually true for Standard Italian:

(301) [ˈvinnə] 'sell!' (IMP) (Lucanian)

(302) [ˈpɔrta] 'carry!' (IMP) (Standard Italian)

However, if we add (clitic) pronouns to these forms, the stress in Lucanian shifts onto the clitics, whereas in Italian it stays on the verb:

(303) [vinnə- ˈmi-lə] 'sell it to me!' (Lucanian)

(304) [ˈpɔrta - me - lo] 'carry it for me!' (Standard Italian)

The difference is that clitics get integrated into the same phonological word as the verb, and then stress gets assigned to the antepenultimate position within that word. In Standard Italian, on the other hand, the phonological word keeps the size of just the verb, and the clitics stay outside (they may form a phonological word in their own right).

One can of course wonder what the explanatory value of the phonological word is precisely. If we set apart prefixes and clitics, we can describe Northern Italian *s* Voicing by referring to the remaining material. But saying that this remainder actually is a constituent called the phonological word and using this *s* voicing as evidence for its existence is circular. We might just as well say that *s* voicing in Italian applies in a morphosyntactic word, but not across the boundary with clitics.

This is indeed a criticism that has sometimes been raised against phonological constituency across the word. A typical answer to this is to show that different kinds of evidence converge on this particular solution. If there are many phonological processes which have a similar restriction, we can see that as an indication that we are onto something.

It is well-known, for instance, that Italian is not alone in keeping prefixes outside of the phonological word. In this respect, prefixes seem to be typologically different from suffixes; we know quite a number of languages in which the former are more separate from the stem than suffixes than we know languages in which this is the other way around, if the former can be convincingly shown to exist at all.

Another example of the relative independence of prefixes we find in German. In this language, we consonants can form the onset of a following vowel across a stem-suffix boundary, but not across the boundary between a prefix and the stem.

For instance, from the noun *Ehr* 'honour' [eːr] one can derive a verb by adding the negative verbalizing prefix *ent*. The resulting word is *entehr* 'dishonour' [ɛnteːr]. The syllabification of this word is [ɛnt.ʔer]: the [t] does not get syllabified with the following [e]; since every German syllable needs to have an onset, this vowel gets a default glottal stop instead.

If we inflect this verb further, for instance with the infinitival ending *-en*, we get *entehren* 'to dishonour' [ɛnteːrən], which is syllabified as [ɛnt.ʔe.rən].

Notice that in this case, the [r] does skip the boundary between stem and affix to syllabify with the initial schwa of the infinitival ending. There thus is indeed a prefix-suffix asymmetry, and it is of the sort we are looking for: the prefix is more independent from the stem than the suffix. This difference is not due to a difference in morphological structure: if anything, the suffix is morphologically 'further away' from the stem than the prefix. Morphologically, we would first derive *entehr* from *Ehr*, and add *en* after this.

As far as we know, there are no languages which work in the opposite way, so that suffixes behave as more independent than prefixes from a phonological point of view even if the arguable are morphosyntactically closer. In itself this is not a complete argument in favour of the prosodic word, if only because prosodic theory does not predict the asymmetry — the opposite could be described as well. But using the notion of a phonological word — similar to but not necessarily always congruent with the morphosyntactic word — at least definitely helps us give a good description of these facts.

Vowel harmony

An important kind of evidence we need to discuss is *vowel harmony*: in quite a lot of languages, we can divide the set of all vowels into at least two subsets, for instance front vowels and back vowels. All vowels in a word are then taken from the same subset. We can describe this in autosegmental terms as saying that the feature [front] spreads within the word, as described in section 4.2.

vowel harmony

The question, obviously is how to describe this 'word', and it is often assumed that the relevant notion of the word (at least in many languages) is indeed prosodic rather than morphological or syntactic.

One concrete example comes from Hungarian. In this language, the suffix meaning 'with' takes on the form *-val* when the preceding stem ends in a back vowel, but the form *-vel* when it ends in a front vowel:

(305) a. Front vowel: *egérrel* 'with mouse', *Ágnessel* 'with Ágnes'
b. Back vowel: *fogóval* 'with pincers'

(There is also some kind of assimilation of the first consonant of *-val* going on, which we ignore.) Notice that stems themselves are not necessarily harmonic: the name Ágnes contains a back vowel followed by a front vowel. Furthermore, in cases such as this, there is actually variation with respect to the choice of the vowel in the suffix. It can also adapt to the first vowel rather than to the second (the second vowel in *Ágnes* can be 'transparent'):

(306) **egérral* 'with mouse', *Ágnessal* 'with Ágnes'

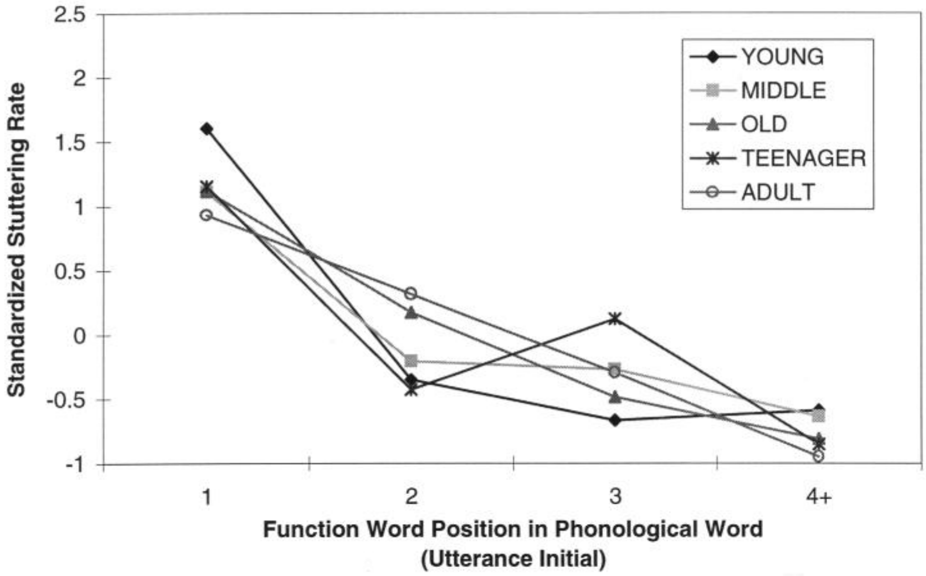
However, things work differently for compounds. If we suffix a compound, the affix can harmonize only to the second part, never to the first part:

(307) **madárlessal* 'with birdwatching', *madárlessel* 'with birdwatching' (from *madar* 'bird' and *less* 'peek')

The morphosyntactic structure of this form is something like [[[madár][les]]sel], but the phonology rather behaves as if it is [madár][[les]sel]. The required phonological structure thus is not precisely congruent with the morphological structure: there are two separate phonological words, but there is one, complex, morphosyntactic word. This is a classical argument for prosodic wordhood.

A final piece of evidence for phonological words come from the phenomenon of stuttering, which most people do sometimes; and some people do so often that it becomes problematic for them. We typically assume that in English functions words (such as determiners *the, a* or the preposition *to*) are grouped in one phonological word together with adjacent lexical words (such as nouns and verbs).

It turns out that people for some reason stutter more on function words than on lexical words, and furthermore that they tend to stutter most if the function word is the first within a phonological word group, as the



(308)

The data are given for different groups of patients; as you can see, for this particular phenomenon it does not really matter whether they are ('young', 'middle' or 'old') children, teenagers or adults: they all show a stronger 'standardized stuttering rate' on a function word when it is on the first position than when it is in the 2nd, 3d or 4th position of the phonological word.

We suspect that the reason for this is *planning* speech. One does not stumble on a function word because that word is itself problematic, but because one is then planning the next (lexical) word. The graph above confirms this idea, but it also shows that apparently the phonological word is a unit of planning: typically, function words at the end of a phonological word may still be followed by a lexical word. But since that lexical word is not in the same phonological word, this apparently does not matter.

8.3 The Phonological Phrase

If we move one level higher up in the prosodic hierarchy, we arrive at the level of the phonological phrase. Like in the case of the phonological word, this constituent has a clear syntactic counterpart: the syntactic phrase (XP, in many theories of syntax). Like in the case of the word, the phonological phrase and the syntactic phrase do not always exactly coincide; otherwise there would of course be no reason to distinguish the two.

A classical argument for non-matching prosodic and syntactic structure comes from English, where arguments have been provided for phonological phrasing of the following type:

(309) *($_{\phi}$this is the cat)($_{\phi}$ that chased the rat)($_{\phi}$ that stole the cheese)*

The ϕ here is the Greek letter Phi, and is the conventional way of abbreviation *phonological phrase*. Phonologically, this sentence consists of three phrases: the positions where I put the dashes are the positions where a speaker could pause, and furthermore there is typically some emphasis on the last word of every phrase.

Yet these three constituents do not necessarily correspond to syntactic constituents. The syntactic structure would be something like this:

(310) [*this is the cat* [*that chased the rat* [*that stole the cheese*]]]

Only the last phrase, *that stole the cheese*, therefore corresponds precisely to a syntactic constituent, but for instance *that chased the rat* does not in any way.

This does not mean, on the other hand, that prosodic and syntactic constituency are completely independent from each other. In particular, every *left* syntactic boundary corresponds to a boundary also in the phonology. However, within the phonology it corresponds both to a left *and* to a right boundary. The reason for this is that phonological constituents do not have *recursion* **recursion** (embedding) in the way in which syntactic constituents do. In the syntactic constituent one sentence ('that stole the cheese') forms an integral part of another sentence ('that chased the rat that stole the cheese').

Phonological phrases do not contain other phonological phrases in the same way. Instead, phonological constituents are usually thought to be restricted by the so called Strict Layer Hypothesis:

(311) **Strict Layer Hypothesis:** Every foot is dominated by a phonological word, every word is dominated by a phonological phrase, every phonological phrase is dominated by an intonational phrase (etc.)

One cannot 'skip layers': feet are never dominated by a phonological phrase directly, and one can also not go back: a word is not dominated by a foot, and not even by a phrase. The phonological structure is organized into layers, which look more or less like autosegmental structure.

Phonological constituents are thus derived from syntactic constituents: typically, given a syntactic structure one can construct the phonological structure reflecting it, but not the other way around: from the fact that (309) has three phrases we cannot conclude which one is contained in which other one in the syntax.

In several Bantu languages, vowel lengthening is an indication for phonological phrases: the penultimate syllable in every phonological phrase lengthens. The following example is from Chichewa, a language which has received a lot of attention in the phrasing literature:

(312) a. *mleéndo* 'visitor'
b. *mlendó uuyu* 'this visitor'

The antepenultimate vowel of *mlendó* is long at the end of a phrase, but not when some word follows it in the same phrase. (Another process that is sensitive to phrase boundaries is tone retraction, moving a tone from a final to a penultimate syllable, but we will ignore this here.)

The verb phrase is (usually) phrased together in Chichewa, as is the subject:

(313) a. *(mwaána)(anaményá nyuúmba)*
child SM-hit house
'The child hit the house'
b. *(mwaána)(anaményá nyumbá ya bwiino))* '
'
child SM-hit house of good
'The child hit the good house'

As you can see, the word *nymba* has a different tonal distribution and different syllable length whether it occurs at the complete end of the sentence (or phrase) or whether some other word follows it.

Another kind of phenomenon that is often taken as evidence in favour of phonological phrases is that some process is restricted to happen within a word. For instance, in Bengali a word-final *r* assimilates (completely) to the first consonant of the next word, but only if this word occurs in exactly the same phonological phrase:

(314) a. (ø^{L}mor^H)(tʃa^L dor^H)(ta^L rake^H)(die^L tʃe^H)
'Amor gave the scarf to Tara'
b. (ø^{L}motʃː a^L dotːa rake^H)(die^L tʃe^H)
'Amor gave the scarf to Tara' (fast speech)

One important thing to notice here is that the prosodic structure is not just dependent on the syntactic structure, as the two examples in (314) are exactly the same in this respect. The only difference is *speech rate*, the number of syllables which are uttered per minute. Prosodic structure can be determined by this: the faster one speaks, the stronger the tendency to put a lot of phonological material in one constituent. This is true most often for prosodic structure above the level of the word.

speech rate

That invoking prosodic phrasing is not just some complicated way of expressing that *r* assimilation affects fast speech more than slow speech is shown by the fact that we can detect phrases also in some other way in Bengali: every phrase starts with a low tone and ends in a high tone. The assimilation and the tone pattern thus converge on exactly the same constituency.

The tones themselves are also worth noticing. Where do they come from? So far, tones had a lexical origin, they were features of a certain word, sometimes floating and at other times lexically prelinked to a certain syllable. But this cannot be the case here, as *any* phonological phrase has the same tonal structure, regardless of the words which are used in it.

This means that the tones belong to the prosodic constituents rather than to the individual words: they are *boundary tones*, linked inherently to the edges of phonological phrases. Boundary tones are one type of *intonation tones*, tones which do not have the function of expressing lexical contrast but rather of making up the tonal melody, helping the listener to parse the stream of sounds in some initial kind of syntactic structure. This may in general be a function of prosodic structure: to guide the listener in figuring out what the syntactic structure is of the sentence he is trying to hear.

boundary tones

intonation tones

8.4 The Intonational Phrase

The highest levels of phonological structure is the Intonational Phrase. (In theory, there is one level that is even higher, that of the *Utterance*, which combines every thing a speaker says within one conversational turn. In practice, very little phonological work has been done on this constituent, and we will ignore it here.)

Utterance

As the name suggests, the *Intonational Phrase* is typically the domain of intonational phenomena, i.e. those that have to do with sentence melody. The size of the Intonational Phrase is typically that of the full sentence (a main clause with all dependent clauses).

Intonational Phrase

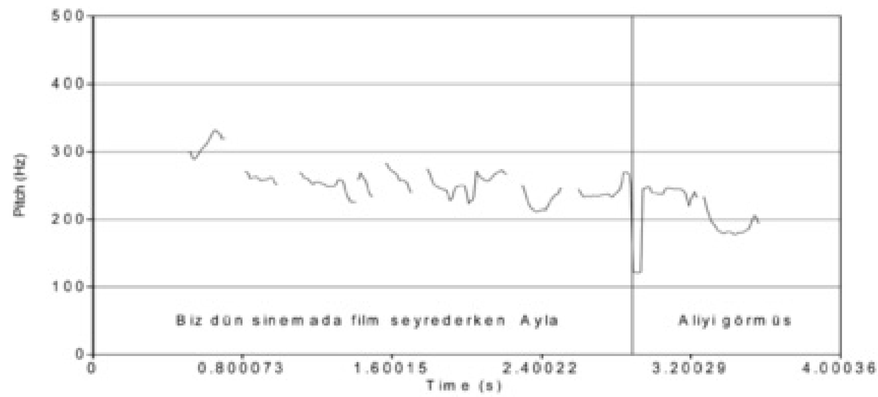
One function of intonation in many languages is to denote sentence types, such as the difference between declarative sentences and questions. Take for instance the following two sentences from Turkish:

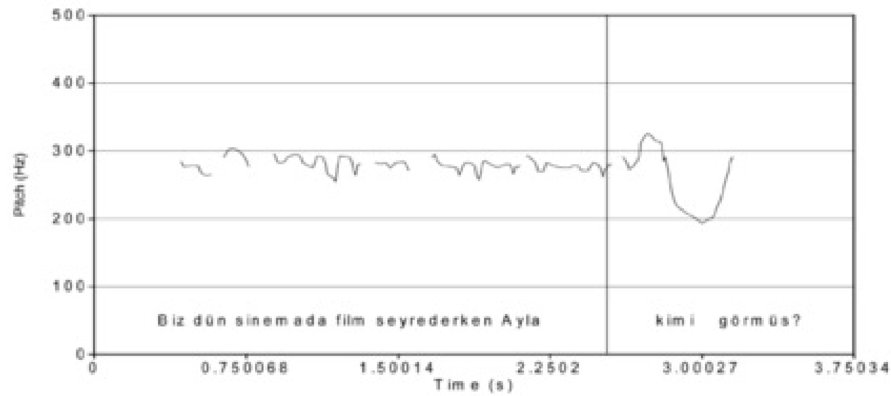
(315) a. *Biz dün sinema-da film seyr-ed-er-ken Ayla Ali-yi gör-mus*
we yesterday cinema-LOC film watch-AUX-AOR-ADV Ayla Ali-ACC see-PERF
'Ayla saw Ali yesterday while we were watching a film at the cinema.'

b. *Biz dün sinema-da film seyr-ed-er-ken Ayla kim-yi gör-mus*
we yesterday cinema-LOC film watch-AUX-AOR-ADV Ayla wbho-ACC see-PERF

'Who did Ayla see yesterday while we were watching a film at the cinema?'

The two sentences have exactly the same word order and almost exactly the same words. Yet one is a question and the other one is not. The way in which this difference is expressed is by intonation: like in many languages, questions end in a relatively high pitch, as you can see in the following pictures:

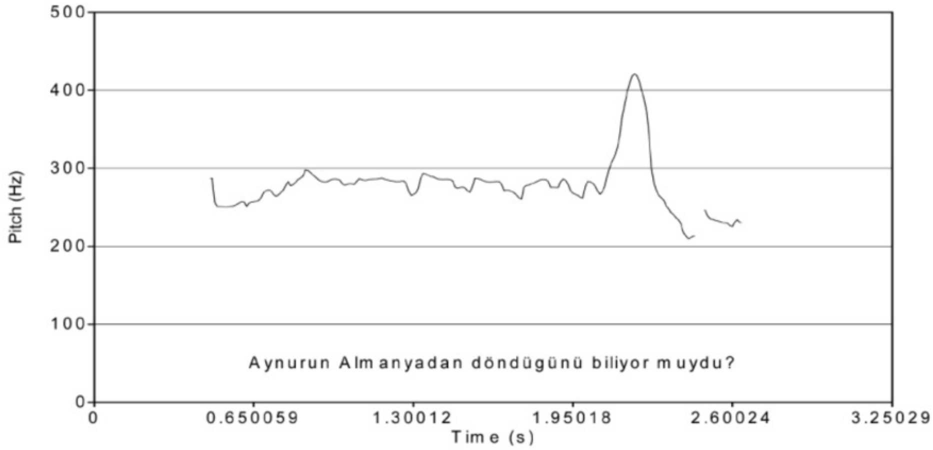
(316) a.


b.


***wh*-question**

This is a so-called wh-*question* containing a question ('*wh*') word, in this case *who*. Yes/no-questions show the same high tone:

(317) a. *Aynur'un Almanya-dan dön-düğ-ün-ü bil-iyor mu-ydu?*
Aynur-GEN Germany-ABL return-COMP-3SG-POSS-ACC know-IMPF Q.PART-P.COP
Did s/he know that Aynur had returned from Germany?



b.

You can also see here that the tone is actually not necessarily attached to the very last syllable, but to the last syllable carrying emphasis or stress. The scholarly literature has converged on assuming that this high pitch is an autosegmental tone which gets linked to the edge of an intonational phrase. This is usually notated as H% — the % marks that this is a so-called *boundary tone*. Thus the *wh*-question above could be represented as follows:

boundary tone

(318) *(($_{IP}$ Biz dün sinema-da film seyr-ed-er-ken Ayla kim-yi gör-mus H%)$_{IP}$*

The boundary tone should be seen as floating, and attaching to the closest vowel, in this case that of the suffix *-mus*. Obviously, English does not have any *lexical tones* which distinguish words from each other the way e.g. the Bantu languages in section 4.1 (p. 65) do. But like Turkish, English does have intonational tones, such as a High boundary tone to indicate questions.

lexical tones

Although intonational patterns are typically the main realisation of intonational phrases, it has sometimes been claimed that also other kinds of phonology are sensitive to the edges of these constituents. For instance, the Tuscan variety of Italian has a process which turns plosives such as /k/ into fricatives such as [h]; however it does so only *inside* intonational phrases, not at the edges:

(319) a. *($_{IP}$Hanno* [h]*atturato sette* [h]*anguri appena nati)$_{IP}$*
'They have captured seven newly born kangaroos.'

b. *($_{IP}$Almeri*[h]*o)$_{IP}$ ($_{IP}$*[k]*uando dorme solo)$_{IP}$ ($_{IP}$*[k]*ade spesso dall'ama*[h]*a)$_{IP}$*
'Almerico, when he sleeps alone, often falls out of the hammock.'

In the latter example, the comma's of the written sentence correspond to intonational boundaries when pronounced — also when reading aloud the English translation you will notice that there are tonal things going on at the edges. But the precise (left) edges of those constituents also seem to prevent /k/ from leniting. This kind of data can be taken as an indication that also fairly 'low level' segmental phonology, referring to individual features such as [continuant], can be sensitive to these 'higher-order' constituents.

8.5 Morphological evidence for the prosodic hierarchy

In the preceding sections, we have seen some 'purely' phonological evidence in favour of the prosodic hierarchy. As pointed out already several times, this evidence is of two types: some phonological process happens only within a phonological constituent of a certain type, or it happens at the boundaries of such a constituent. This is interesting evidence, but it has been pointed out as well that it is mostly evidence for the *boundaries*: processes either require those to be absent (the first type) or present (the second type).

We can find more direct evidence for the constituency of phonological objects we have to turn to the interaction of phonology with word formation procesεses, morphology. In previous chapters we have largely discussed phonology as a world on its own, but there are many indications that phonology interacts with the way in which words are formed in a number of ways.

We will discuss this interaction in more detail in the next chapter, but here we concentrate on the specific phenomenon of *Prosodic Morphology*, types of word formation which refers to elements of prosodic structure, such as morae, syllables, feet and phonological words. (Higher-order prosodic constituents play less of a role, because phonological and intonational phrases are typically too big to match morphosyntactic words.) Examples of prosodic morphology are infixation and reduplication; these processes are rather rare in English, but in some other languages they abound. One basic claim of the theory of prosodic porphology is (McCarthy and Prince, 1998):

(320) Morphological processes that refer to phonological structure use the same prosodic structures as ordinary phonology: morae (μ), syllables (σ), feet (Ft) and phonological words (ω).

expletive infixation

Although English does not have a lot of prosodic phonology, it does have one process: so-called *expletive infixation*. In some varieties, one can insert an expletive such as *bloody* — or forms which are even more taboo — within a word to give some special effect. However, famously, it is not possible to do this at just any position in the word:

(321) a. *fan-bloody-tastic*
b. **fa-bloody-ntastic*
c. **fantas-bloody-tic*

Speakers have quite clear intuitions about what is and what is not possible. In tests where they are asked to apply expletive infixation to new terms, they will do so without a lot of variation:

(322) a. *amalga-bloody-mated*
b. **amalgam-bloody-ated*
c. **amal-bloody-gamated*

We can understand these judgements as a wish to keep the prosodic constituency of the original word intact as much as possible. In (321b) and (322b), we have inserted the expletive within a syllable (*fan* and *ma*), and the result of this is bad. In (321c) and (322c), we did a similar thing to a phonological foot, which in English is always a trochee (*tastic* and *malga* respectively), which apparently is also wrong.

We have to break open at least the phonological word in expletive infixation, because otherwise there would be no infix; but speakers of English intuitively seem to feel that other prosodic constituents should be respected. This in itself is evidence that these constituents somehow are part of their knowledge of language.

This is the kind of intuitions we find attested in prosodic morphology: speakers show knowledge of the prosodic hierarchy in creating new words. One process which is quite widespread and in which this happens is *reduplication*. In this process, part of a word is doubled to reveal some special meaning. **reduplication** The following is an example from Ilokano. As in many languages, reduplication expresses plurality on nouns in this language:

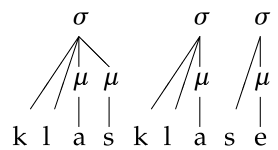
(323)	káldíŋ	'goat'	*kál*-káldíŋ	'goats'
	púsa	'cat'	*pús*-pusa	'cats'
	kláse	'class'	*klas*-kláse	'classes'
	jyánitor	'janitor'	*jyan*-jy'anitor	'janitors'
	róʔot	'litter'	*roː*-róʔot	'litter (pl.)'
	trák	'truck'	*traː*-trák	'trucks'

In these examples, the copied part (the *reduplicant*) has been italicized. It is **reduplicant** easy to see that this reduplicant is a heavy syllable $\sigma_{\mu\mu}$ in all cases. The idea is that the plural suffix in Ilokano takes this shape: it is an 'empty' heavy syllable which has to be materially filled with segments from the base. Note, by the way, that the mora theory is a convenient way to express this.

We do not just copy the first syllable of the stem. The first syllable of *pusa* presumably is *pu*. In order to fill the heavy syllable template, however, we have to add the *s* which is part of the second syllable. The reason why the vowel is lengthened in *roː-róʔot* and *traː-trák* has something to do with preferences of syllable structure, which we will not discuss here.

We thus imagine the derivation of a reduplicated plural in Ilokano in the following way:

(324)

Input	*Output*
$\sigma_{\mu\mu}$ 'plural' + klase 'class'	

Languages can also choose to specify light syllables (σ_{μ}) as the reduplicant, and as a matter of fact Ilokano provides an instance of this as well, in a suffix which meens 'covered with':

(325)	buneŋ	'buneng'	si-*bu*-buneŋ	'carrying a buneng'
	jyaket	'jacket'	si-*jya*-jyaket	'wearing a jacket'
	pandiliŋ	'skirt'	si-*pa*-pandiliŋ	'wearing a skirt'

Again, this is not just a process of copying a syllable, witness the last example: we do not copy all material from the heavy syllable *pan* in the stem, but only just enough to fill the light syllable template.

Higher-order structure can also function as a template for reduplication. In Diyari we copy a Foot to derive various morphological effects:

(326)	wil̪a	*wil̪a*-wil̪a	'woman'
	kanku	*kanku*-kanku	'boy'
	kulkuŋa	*kulku*-kulkuŋa	'to jump'
	tʲiplarku	*tʲilpa*-tʲiplarku	'bird species'
	ŋankan̪t̪i	ŋ*anka*-ŋankan̪t̪i	'catfish'

Again, we see that it is not just the first two syllables (or the first foot) of the word which are copied: the final syllable of the reduplicant is always open, even if the second syllable of the base is not.

Reduplication can also sometimes be total: we then copy the whole phonological word. Indonesian plural formation is a case in point:

(327)	wanita	woman	wanita-wanita	women
	mašakarat	society	mašakarat-mašakarat	societies

inchoative meaning

It will come as no surprise that the phonological word is not always exactly the same as the morphosyntactic word, and in particular that prefixes are not always copied together with the stem. In the Bantu language Kihehe, we find a reduplication process that has an *inchoative meaning* — it denotes the start of an event:

(328)	kú-haáta	'to ferment'	kú-haáta-haáta	'to start fermenting'
	kú-gohomóla	'to cough'	kú-gohomóla-gohomóla	'to start coughing'

Everything is copied, including the stem and the ending *-a* in these words; but the prefix is not. We can thus say that Kihehe reduplicates the phonological word; and that such words are formed in a way which is parallel to that of Italian.

Axininca Campa gives another example of a language in which we copy a phonological word, but with a certain restriction. In this language, we find the following pattern:

(329)	kawosi-*kawosi*	'bathe'
	koma-*koma*	'paddle'
	osampi-*sampi*	'ask'
	osaŋkina-*saŋkina*	'write'

If the stem starts with a consonant, it is completely reduplicated; if it starts with a vowel, it is reduplicated except for the initial vowel. The reason for this presumably is syllable structure: by not copying the initial vowel, we avoid 'unnecessary' violations of the constraint ONSET, requiring every syllable to start with an onset consonant.

(330) REDUPLICATEMAX (RM): Reduplicate all material from the stem.

(331) a.

/osampi/+RED	NODELETION	ONSET	RM
o.sam.pi.o.sam.pi		**!	
☞o.sam.pi.sam.pi		*	*
sam.pi.sam.pi	*!	*	

b.

/kawosi/+RED	NODELETION	ONSET	RM
☞ka.wo.si.ka.wo.si			
ka.wo.si.wo.si			*!*

Emergence of the Unmarked

We see a very important effect here, which distinguishes OT in a favourable way from parametric theories. In a theory of the latter type, we would need to say that the ONSET parameter is set 'on' in Axininca Campa: witness words such as *osampi*, the language allows onsetless syllables. But then we cannot explain why we all of a sudden find a restriction on them in reduplicated forms.

In OT, the situation is different: the constraint ONSET is sufficiently low-ranked — below the relevant faithfulness constraint — to make its effect invisible in ordinary words. But in reduplicated words, faithfulness is no longer important (the first vowel of *o* is present anyway), so that now all of a sudden we can see the universal constraint ONSET can be seen at work. This is called an effect of *the emergence of the unmarked* (*TETU*), and it is at present the strongest argument in favour of OT over parametric theories. TETU

TETU effects abound in reduplicative systems. For instance, Sanskrit usually allows complex onsets, but when these onsets are reduplicated (in the perfective forms of verbs), they are simplified. Reduplication in this case is to a word.

(332) a. *pa*-prat^h-a 'spread'
b. *ma*-mna:-u 'note'
c. *sa*-swar 'sound'
d. *da*-d^{h}wans-a 'scatter'

The constraint which shows up here in the reduplicated form is the one against complex onsets:

(333) a. RED=σ_μ: The reduplicative suffix is a monomoraic syllable.
b.

RED + /prat^h/	NODELETION	*COMPLEX	RED=σ_μ	RM
pra.pra.t^{h}a		**!		t^{h}a
☞pa.pra.t^{h}a		*		rt^{h}a
pa.pa.t^{h}a	*!			t^{h}a
pa.t^{h}a.pra.t^{h}a		*	*!	

I introduced a small piece of new notation in this tableau: in the last column, I list the segments which violate the constraint. I could just as well have given one asterisk for every segment, but this notation gives slightly more insight into what is actually going on.

We also find TETU effects at the level of segmental structure, for instance in Akan (the facts have been slightly simplified):

(334) a. siʔ → *si*.siʔ 'stand'
b. seʔ → *si*.seʔ 'say'
c. buʔ → *bu*.buʔ 'bend'
d. soʔ → *su*.soʔ 'seize'

This pattern looks very much from what we have seen for reduction. In terms of Element Theory we could state the following markedness and faithfulness constraints:

(335) a. NOCOMPLEXVOWEL: Only allow primary vowels (markedness)
b. KEEPA: Don't delete the element A

These constraints again interact to get a TETU effect:

(336)

RED + /seʔ/	KEEPA	NOCOMPLEXVOWEL	RM
se.seʔ		**!	
☞si.seʔ		*	|A|
si.siʔ	*!		

Infixation and shape restrictions

infixation

A second well-known example of a prosodic morphological process next to reduplication is *infixation*, the positioning of an affix not at the left (prefix) or right (suffix) edge. Consider the placement of the third person singular masculine possessive suffix *-ka* in Ulwa (a language from Nicaragua):

(337)

bas	'hair'	*bás-ka*	'his hair'
ki:	'stone'	*kí*:*-ka*	'his stone'
su:*lu*	'dog'	*sú*:*-ka-lu*	'his dog'
asna	'clothes'	*as-ka-na*	'his clothes'
sana	'deer'	*sana-ka*	'his deer'
amak	'bee'	*amak-na*	'his bee'
sapa:	'forehead'	*sapa*:*-ka*	'his forehead'
siwanak	'root'	*siwa-ka-nak*	'his root'
ana:*la*:*ka*	'chin'	*ana*:*-ka-la*:*ka*	'his chin'

At first sight, it looks as if *-ka* sometimes behaves as a suffix, but sometimes it is also inserted inside the word. On closer inspection, the generalisation is that *-ka* comes after the first syllable of the word, if that syllable is heavy, and otherwise it comes after the second syllable. An insightful way to see this, is to say that the possessive behaves as a suffix to the first (iambic) foot of the word.

The hypothesis of Prosodic Morphology is that infixation is always of this type: it never means putting an affix just in some random position inside the stem: it is always prefixed or suffixed to a prosodic constituent. This type of analysis is also often used to show the advantages of Optimality Theory. Look at the following examples from Tagalog with the infix *-um*:

(338) *um-alis* 'leave'
t-um-awag 'call PERF. ACTOR TRIGGER
gr-um-adwet 'graduate'

In this case, *um* sometimes looks like a genuine prefix, and sometimes it looks like an infix; the generalisation is that it is prefixed if the word starts with a vowel and infixed otherwise. Within OT, we can give an elegant description of these facts: by infixing, we prevent an unnecessary violation of the constraint NOCODA. We do this at the cost of violating a (new) instance of an Alignment constraint, one forcing the left edge of the affix to be aligned to the left edge of the word; in other words, making the affix to behave like a real *pre*fix.

(339) a. ALIGN(*-um-*, L, ω, L): The left edge of *-um-* should correspond to the left edge of the word (count violations in segments).

b.

/um+tawag/	NODELETION	NOCODA	ALIGN
um.ta.wag		**!	
☞tu.ma.wag		*	*
ta.wu.mag		*	**!*
u.ta.wa	*!*		

c.

/um+alis/	NODELETION	NOCODA	ALIGN
☞u.ma.lis		*	
a.um.lis		**!	*
a.lu.mis		*	*!*
u.ma.li	*!		

Notice that it follows from the principles of the theory that there is a relation between the shape of the infix — *VC* — and its infixal behaviour. We predict that there could not be a language where an affix *mu* could display the same behaviour:

(340) Non-existing language:
a. mu+alis → a.mu.lis
b. mu+tawag → mu.ta.wag

The reason is that infixation in this case does not help:

(341)

/mu+alis/	NOINSERTION	ONSET	ALIGN
☞mu.a.lis		*	
a.mu.lis		*	*!

No matter where we place the infix, there will always be a violation of the constraint ONSET; and this hypothetical language will allow onsetless violations, since it has hypothetical words of the shape *alis*.

Infixation and reduplication are sometimes combined. For instance, in Samoan the σ_μ reduplicant is prefixed to the stress foot:

(342) a. fa.náu → fa.*na*.náu 'be born'
b. a.lófa → a.*lo*.lófa 'love'

The following paradigm (from Timugon Murut) is also of interest in this connection:

(343) a. bulud → *bu*-bulud 'hill/ridge'
b. dondo → *do*-dondo? 'one'
c. indimo → in-*di*-dimo 'five times'
d. ompod → om-*po*-pod 'flatter'

This combines the two types of prosodic morphology we have seen so far: infixation and reduplication. The affix clearly reduplicates a light syllable of the stem; but in some cases (if the stem starts with a vowel) it also is infixed inside the stem, so that it does not reduplicate the first syllable, but the second one.

This example is of interest, since it seems to violate a generalisation we just made: that there are no phonological infixes of the shape CV. The reason is that in cases of infixed reduplication we *do* avoid unnecessary violations of the constraint ONSET. If we would copy the first syllable, we would create an 'unnecessary' onsetless syllable, which can be avoided if we copy the second one instead. Still, we would like the infix to be as much to the left — as much as a prefix — as possible:

(344) a.

/ompod/+RED	NOINSERTION	ONSET	ALIGN
om.om.pod		**!	
☞om.po.pod		*	om

b.

/bulud/+RED	NOINSERTION	ONSET	ALIGN
☞bu.bu.lud			
bu.lu.lud			bu!

Hypercoristics formation

A third process of prosodic morphology next to reduplication and infixation is nick-name (*hypocoristics*) formation. In many languages, shorter versions can be used of personal names, for instance to express affection. In these cases, the hypocoristics assume the shape of some well-described prosodic constituent:

hypocoristics

(345)

Name	Hypocoristic	
ti	tiičan	
šuusuke	šuu-čan	*šuusu-čan
yoosuke	yoo–čan	*yoosu-čan
taizoo	tai-čan	*taizo-čan
kinsuke	kin-čan	*kinsu-čan
midori	mii-čan	*mi-čan
	mit-čan	
	mido-čan	
wasburoo	waa-čan	*wa-čan
	wasa-čan	
	sabu-čan	
	wasaburo-čan	*wasabu-čan

The Japanese hypocoristic consists of a shortened version of the original name plus the suffix *-čan*. These examples show is that not just any shortening will do; we can observe that all the correct hypocoristics consist of an even number of morae, whereas the wrong versions all have an odd number of morae. In terms of the typology of stress feet from the previous class, this implies that the base to which *-čan* is attached will consist of a number of moraic trochees.

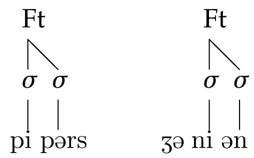
We can observe the restriction to bases of a certain shape also outside the domain of hypocoristic formation or prosodic morphology proper. For instance, Dutch has two productive plural suffixes, *-ən* and *-s*. The first one is generally chosen after stems ending in an unstressed syllable, and the second one after a stressed syllable (underlining marks stress):

(346) a. *genie* 'genius' *genieën* 'geniuses'
fabriek 'factory' *fabrieken* 'factories'

b. *familie* 'family' *families* 'families'
pieper 'potato' *piepers* 'potatoes'

Why do we observe this distributional paradigm? Notice that because of this effect, plural words tend to end in a (syllabic) trochee: a stressed syllable followed by an unstressed syllable.

(347)



There is an importance difference between the Japanese hypocoristics suffix *-čan* and the Dutch plural suffix: the former requires its *input* to take a specific trochaic shape, whereas the latter makes sure that the *output* has this particular shape. Both of them are in support of the claim — which typological study seems to have confirmed — is that when morphology requires morphemes or words to have refer some specific shape, such shapes are always taken from the stock of prosodic phonology. To some extent, this is of course not surprising: phonological

8.6 Exercises

1. 'Prosodic structure is derived from syntactic structure, but not the other way around.' Explain.
2. 'The phonological structure is organized into layers, which look more or less like autosegmental structure.' (p. **??**) Explain.
3. The Australian language Yidiɲ has a process lengthening the antepenultimate syllable of words with an odd number of syllables. This process also applies to suffixed words:

   (348) a. *gudagagu* 'dog-PURP'
   b. *muɖam* 'mother-ABS'
   c. *muɖa:mga* 'mother-PURP'

   However, some clusters of suffixes behave like independent units for this process:

   (349) *guma:ri daga:ɲu* 'to have become red' (from /gumari/ 'red', /daga/ (INCHOATIVE) and /ɲu/ (PAST)

   a. Construct an argument in favour of the phonological word from this example.
   b. Forms which behave like (349), typically include bisyllabic suffixes. Can you think of a reason why?
4. Korean has a process of intersonorant voicing of obstruents (so a voiceless obstruent becomes voiced between two vowels or sonorants). Study the following examples, and decide what is the prosodic domain of voicing, and how this (roughly) corresponds to morphosyntactic structure.

   (350) a. /apəci/ [abəji] 'father'
   b. /kı cip/ [kıɟip] 'that house'
   c. /motın kırim/ [modın gırim] 'every picture'
   d. /ʃuni-ıj cip/ [ʃunoıj ɟip] 'Suni's house'
   e. /kırim-ıl pota/ [kırim-ıl boda] 'look at the picture'
   f. /kæka canta/ [kæga canda] 'de dog is sleeping' (lit. 'dog sleeps')
   g. /horaŋi-wa kojaŋi/ [horaŋi-wa kojaŋi] 'the tiger and the cat'

5. In Kinande, verbs can be reduplicated to mean (for instance) that the action that is described is done little by little. In the following, you see how the reduplication works for stems of different shapes:

   (351)

a.	*Consonant-initial*	huma	huma-huma	'beat'
		humira	huma-humira	'beat for'
		humirana	huma-humirana	'
b.	*Vowel-initial*	esa	eses-esa	'play'
		oha	ohoh-oha	'pick'

c. *Monosyllabic*	swa	swaswa-swa	'grind'
	ta	tata-ta	'bury'

6. Order the constraints in (300) in such a way that you can derive the different phonological forms of (298). You may need a special constraint for the compound. Can you think of a form for this constraint?
7. Here are a few examples of reduplication in Mokilese. What is the size of the reduplicant?

(352)

	reduplicated	
pɔdɔk	pɔdpɔdɔk	'plant'
mwiŋe	mwiŋmwiŋe	'eat'
kasɔ	kaskasɔ	'throw'
poki	pokpoki	'beat'

8. Look at the following examples of Spanish names and corresponding hypocoristics.

(353)

Name	Hypocoristic
ɑlejandro	ɑle
ɑsunción	ɑsun
ɪsabel	ɪsa
ȷeronimo	ȷero
ɾodrigo	ɾodri

What is the template for this process of hypocoristic formation?

9. Example (328) shows a few instances of reduplication in Kihehe. Here is a further example:

(354) /kú-ita/ [kwiːta] 'to pour' kwiːta-kwiːta 'to pour a little'

As you can see, in this case the prefix *is* copied with the stem. What could be the reason? Give a simple OT analysis; you can freely make up the relevant constraints.

10. Try to reanalyse the examples from English and Bengali in section 8.3 without referring to phonological phrases, but only to syntactic constituents. Does the difference give you a reason to prefer one analysis over the other?
11. The examples below give the beginning and the end for a sign meaning 'shop' in Israeli Sign language and for the beginning and the end for a sign meaning 'shop there' in the same language. Discuss how such data could be used to study phonological word structure in (Israeli) sign language.

(355) a.  

b. 

12. In English, clusters of obstruents tend to assimilate in voicing, either progressively (356a) or regressively (356b). However, this happens only to obstruents within the same word, not across word boundaries (356c).

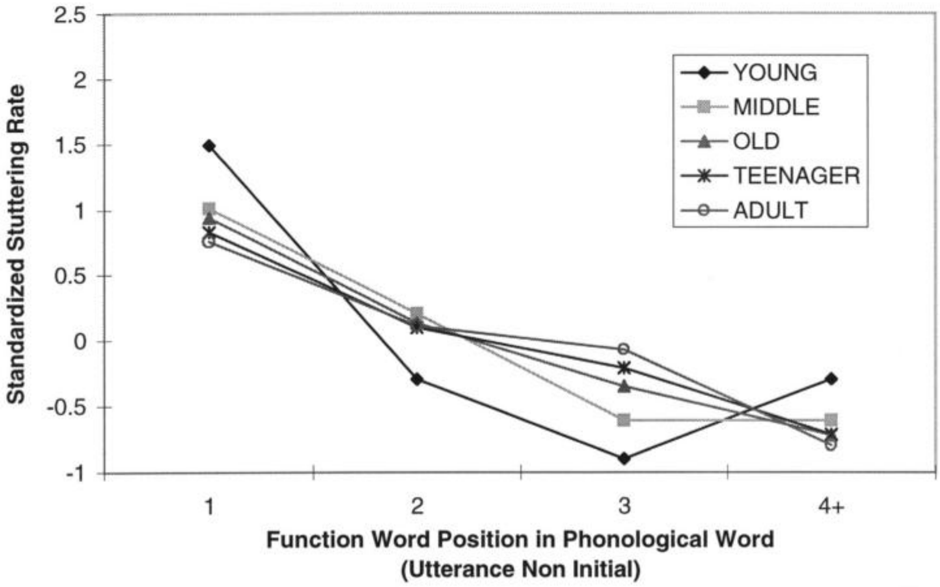
(356) a. *twelve* [twɛlv] + *th* [θ] → *twelfth* twElfθ

b. *cat* [kæt] + *s* [z] → *cats* [kæts]

c. *The cat zooms* [kæt] + *zooms* [zums] → [tz] (*[ts], *[dz])

If you find a sequence of two obstruents of which one is voiced and the other voiceless in English, you can be sure there is a word boundary between them. Is there any reason to think this is a *phonological* word boundary?

13. Example (308) (p. 168) shows a graph of the distribution of stuttering in function words in different positions within *the first phonological word* of the utterance. The following shows a graph from the same study on the same distrbution but now in later phonological words.

(357) 

As you can see, the two patterns are not very different. What does this mean for our evidence with respect to planning and prosodic structure?

8.7 Sources and further reading

Section 1.1. The classical book about prosodic phonology is In recent years, the idea that phonological structures are not recursive has come under attack; see

The paper on stuttering and phonological words from which the figures in (328) and (328) are taken is

Bibliography

Adger, D. (2003). *Core Syntax*. Oxford: Oxford University Press.

Alidou, O. (1997). *A linguistic study of Nigerian language games*. Ph. D. thesis, Indiana University.

Anderson, C. (2018). *Essentials of Linguistics*. Vancouver, BC: BCCampus.

Anderson, J. and C. Ewen (1987). *Principles of Dependency Phonology*. Cambridge: Cambridge University Press.

Anderson, S. R. (1985a). *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations*. Chicago: University of Chicago Press.

Anderson, S. R. (1985b). *Phonology in the Twentieth Century. Theories of Rules and Theories of Representations*. Chicago and London: The University of Chicago Press.

Ashby, M. and J. Maidment (2005). *Introducing Phonetic Science*. Cambridge, UK: Cambridge University Press.

Backley, P. (2011). *An Introduction to Element Theory*. Edinburgh: Edinburgh University Press.

Bakovic, E. (2000). *Harmony, Dominance and Control*. Ph. D. thesis, Rutgers.

Bat-El, O. (2011). Semitic templates. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell Companion to Phonology*, Chapter 108, pp. 2586–2608. Oxford & Cambridge, Mass.: Wiley-Blackwell.

Blake, B. J. (1969). *The Kalkatungi Language*. Australian aboriginal studies. Acton: Australian Institute of Aboriginal Studies.

Bolton, T. (1894). Rhythm. *American Journal of Psychology 6*, 145–238.

Bonatti, L., M. Pena, M. Nespor, and J. Mehler (2005). The role of consonants and vowels in continuous speech processing. *Psychological Science*, 91–103.

Booij, G. (2007). *The Grammar of Words. An Introduction to Morphology*. Oxford: Oxford University Press.

Botma, B. (2004). *Phonological Aspects of Nasality. An Element-Based Dependency Approach*. Ph. D. thesis, University of Amsterdam.

Brockhaus, W. (1995). *Final Devoicing in the Phonology of German*. Max Niemeyer Verlag.

Buckley, E. (2011). Metathesis. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell companion to phonology*, Volume 3. Oxford: Wiley-Blackwell.

Campbell, A. (1959). *Old English Grammar*. Oxford, UK: Clarendon Press.

Carreiras, M. and C. J. Price (2008). Brain activation for consonants and vowels. *Cerebral Cortex 18*, 1727–1735.

Charette, M. (1991). *Conditions on Phonological Government*, Volume 58 of *Cambridge Studies in Linguistics*. Cambridge: Cambridge University Press.

Chierchia, G. (1986). Length, syllabification and the phonological cycle in italian. *Journal of Italian Linguistics 8*(5-33).

Chin, S. B. (2006). Realization of complex onsets by pediatric users of cochlear implants. *Clinilcal Linguistics and Phonetics 20*(7-8), 501–508.

Chomsky, N. (1968). *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.

Chomsky, N. and M. Halle (1968). *Sound Pattern of English*. Harper and Row.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. In J. Kingston and M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, pp. 283–333. Cambridge, UK: Cambridge University Press.

Clements, G. N. and B. Hume (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *The Handbook of Phonological Theory*, pp. 245–306. Cambridge: Blackwell Publishers.

Clements, N. and E. Sezer (1982). Vowel and consonant disharmony in turkish. In H. van der Hulst and N. Smith (Eds.), *The Structure of Phonological Representations*, pp. 213–255. Dordrecht: Foris.

Cole, J. and J. I. Hualde (2011). Underlying representations. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell Companion to Phonology*, Volume 1. Oxford: Wiley-Blackwell.

Conklin, H. (1956). Tagalog speech disguise. *Language 32*, 136–139.

Coulmas, F. (2003). *Writing systems. An introduction to their linguistic analysis*. Cambridge, UK: Cambridge University Press.

Crosswhite, K. (2001). *Vowel Reduction in Optimality Theory*. New York: Routledge.

Cutler, A., J. Mehler, D. Norris, and J. Seguí(1986). The syllable's differing role in the segmentation of french and english. *Journal of Memory and Language 25*, 385–400.

Davis, S. (1990). Italian onset structure and the distribution of *il* and *lo*. *Linguistics 28*, 43–55.

Davis, S. and S.-H. Shin (1999). The syllable contact constraint in korean. an optimality-theoretic analysis. *Journal of East Asian Linguistics 8*, 285–312.

de Lacy, P. (Ed.) (2007). *The Cambridge Handbook of Phonology*. Cambridge, UK: Cambridge University Press.

Dresher, B. E. (2009). *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.

Duanmu, S. (2009). *Syllable structure. The limits of variation*. Oxford, UK: Oxford University Press.

Ernestus, M. and H. Baayen (to appear). The functionality of incomplete neutralization in dutch. the case of past-tense formation. In *Papers in Laboratory Phonology 8*.

Ewen, C. and H. van der Hulst (2001). *The Phonological Structure of Words. An Introduction*. Cambridge: Cambridge University Press.

Fikkert, P. (1994). *On the acquisition of prosodic structure*. Ph. D. thesis, University of Leiden.

Fromkin, V. (Ed.) (1973). *Speech errors as linguistic evidence*. Mouton.

Fromkin, V. A., R. Rodman, and N. Hyams (2018). *An Introduction to Language* (11 ed.). Andover: Cengage.

Gahl, S. (2008). *Time* and *thyme* are not homophones: The effect of lemma frequency on word durations in spontaneous speech. *Language 84*(3), 474–496.

Goldsmith, J. (1990). *Autosegmental and Metrical Phonology*. Oxford: Basil Blackwell.

Goldsmith, J., J. Riggle, and A. C. Yu (2014). *The handbook of phonological theory*. Oxford: Wiley-Blackwell.

Gouskova, M. (2003). Falling sonority onsets, loanwords, and syllable contact. *CLS 37*.

Gussenhoven, C. and H. Jacobs (1998). *Understanding phonology*. London: Arnold.

Gussenhoven, C. and H. Jacobs (2005). *Understanding Phonology*. London: Hodder Arnold.

Gussman, E. (2002). *Phonology. Analysis and Theory*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Hall, N. (2011). Vowel epenthesis. In M. van Oostendorp, C. J. Ewen, E. Hume, and Keren (Eds.), *The Blackwell Companion to Phonology*, Volume 2 of *1576–1596*. Wiley-Blackwell.

Halle, M. and G. N. Clements (1983). *Problem Book in Phonology: A Workbook for Introductory Courses in Linguistics and in Modern Phonology*. Cambridge, Mass.: MIT Press.

Hammond, M. (1999). *The Phonology of English*. Oxford: Oxford University Press.

Harris, J. (1994). *English Sound Structure*. London: Blackwell.

Harris, J. (2005). Vowel reduction as information loss. In P. Carr, J. Durand, and C. Ewen (Eds.), *Headhood, Elements, Specification Headhood, Elements, Specification and Contrastivity. Phonological Papers in Honour of John Anderson*, pp. 119–132. Amsterdam: John Benjamins.

Harris, J. (2011). Deletion. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell Companion to Phonology*, Volume 2, pp. 1597–1621. Oxford and Cambridge, Mass.: Wiley-Blackwell.

Harris, J. and G. Lindsey (1995). The elements of phonological representation. In J. Durand and F. Katamba (Eds.), *Frontiers of Phonology*, pp. 34–79. London and New York: Longman.

Haugen, J. D. (2003). Allomorphy in yaqui reduplicatio. In L. M. Barragan and J. D. Haugen (Eds.), *Studies in Uto-Aztecan*, Volume 5, pp. 75–103. MIT Working Papers on Endangered and Less Familiar Languages.

Hayes, B. (1987). A revised parametric metrical theory. *Northeastern Linguistics Society 17*, 274–289.

Hayes, B. (1989). Compensatory lengthening in moraic phonology. *Linguistic Inquiry 20*.

Hayes, B. (1995). *Metrical stress theory. Principles and Case Studies*. University of Chicago Press.

Hayes, B. (2009). *Introductory phonology*. Malden, Mass., and Oxford, UK: Wiley-Blackwell.

Hoffmann, C. (1963). *A grammar of the Margi language*. London: Oxford University Press.

Hogg, R. (1992). *A Grammar of Old English*, Volume !: Phonology. Oxford, UK and Cambridge, Mass.: Blackwell.

Holt, D. E. (2004). Optimization of syllable contact in old spanish via the sporadic sound change metathesis. *Probus 16*(1), 43–61.

Hulst, H. v. and J. v. Weijer (1995). Vowel harmony. In J. Goldsmith (Ed.), *The Handbook of phonological theory*, pp. 495–534. Oxford: Basil Blackwell.

Humbert, H. (1995). *Phonological Segments. Their Structure and Behaviour*. Ph. D. thesis, University of Amsterdam.

Hume, E. (1998). Metathesis in phonological theory: The case of Leti. *Lingua 104*, 147–186.

Hyman, L. (2008). Universals in phonology. *The Linguistic Review 25*, 85–137.

Hyman, L. (2018). Why underlying representations. *Journal of Linguistics 54*(3), 591–610.

Itô, J. (1986). *Syllable theory in prosodic phonology*. Ph. D. thesis, University of Massachusetts at Amherst.

Ito, J. (1986). *Syllable Theory in Prosodic Phonology*. Ph. D. thesis, University of Massachusetts, Amherst. Published 1988. Outstanding Dissertations in Linguistics series. New York: Garland.

Itô, J. and A. Mester (2003). *Japanese Morphophonemics. Markedness and Word Structure*. Cambridge, Mass.: MIT Press.

Kager, R. (1999). *Optimality Theory*. Cambridge: Cambridge University Press.

Kaye, J. (1989). *Phonology. A cognitive view*. Cognitive Science Series. Tutorial Essays. Hillsdale, New Jersey and Hove and London: Lawrence Erlbaum Asociates.

Kaye, J. (1992). Do you believe in magic? The story of s+C sequences. *SOAS Working Papers in Linguistics and Phonetics 2*, 293–313.

Kaye, J., J. Lowenstamm, and J.-R. Vergnaud (1985). The internal structure of phonological elements. a theory of charm and government. *Phonology Yearbook 2*, 305–328.

Kaye, J., J. Lowenstamm, and J.-R. Vergnaud (1990). Constituent structure and government in phonology. *Phonology 7*, 193–232.

Kennedy, R. (2011). Reduplication in sanskrit. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell Companion to Phonology*, Volume 5, Chapter 119, pp. 2855–2878. Oxford, UK: Wiley-Blackwell.

Kennedy, R. (2018). *Phonology. A Coursebook*. Cambridge University Press.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Oxford, UK: Blackwell.

Kenstowicz, M. (2007). Salience and similarity in loanword adaptation. a case study from fijian. *Language Sciences 29*, 316–340.

Kenstowicz, M. and C. Kisseberth (1979). *Generative Phonology: Description and Theory*. New York: Academic Press.

Keyser, S. J. and P. Kiparsky (1984). Syllable structure in finnish phonology. In M. Aronoff and R. Oehrle (Eds.), *Language Sound Structure. Studies in Phonology Presented to Morris Halle by his Teacher and Students*, pp. 7–31. Cambridge, MA: MIT Press.

Keyser, S. J. and K. N. Stevens (2006). Enhancement and overlap in the speech chain. *Language 82*, 33–63.

Kiparsky, P. (2003). Accent, syllable structure and morphology in ancient greek. In E. M. Athanasopoulou (Ed.), *Selected papers from the 15th international symposium on theoretical and applied linguistics*. ISTAL.

Knight, R.-A. (2012). *Phonetics. A Coursebook*. Cambridge, UK: Cambridge University Press.

Kornfilt, J. (1977). *Turkish*. New York: Routledge.

Krämer, M. (2009). *The Phonology of Italian*. The Phonology of the World's Languages. Oxford: Oxford University Press.

Lahiri, A. and H. Reetz (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics 38*, 44–59.
Lass, R. (1976). *English phonology and phonological theory*. Cambridge: Cambridge University Press.
Levelt, C. C. and M. van Oostendorp (2007). Feature co-occurrence constraints in l1 acquisition. In B. Los and M. van Koppen (Eds.), *Linguistics in the Netherlands 2007*, pp. 162–172. Amsterdam and Philadelphia: John Benjamins.
Liberman, I. Y., D. P. Shankweiler, F. W. Fischer, and B. Carter (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology 18*, 201–212.
Lohmann, A. (2018). *Time* and *thyme* are not homophones: A closer look at gahl's work on the lemma frequency, including a reanalysis. *Language 94*(2), e180–2190.
Lynch, J. (1974). *Lenakel Phonology*. Ph. D. thesis, University of Hawaii. Also, University of Hawaii Working Papers in Linguistics vol. 7.1.
MacKay, D. G. (1972). The structure of words and syllables. evidence from errors in speech. *Cognitive Psychology 3*, 210–227.
McCarthy, J. and A. Prince (1993a). Prosodic morphology i. constraint interaction and satisfaction. Available on ROA 485-1201.
McCarthy, J. and A. Prince (1998). Prosodic morphology. In A. Spencer and A. M. Zwicky (Eds.), *The Handbook of Morphology*, pp. 283–305. Oxford: Blackwell.
McCarthy, J. J. (1979). *Formal problems in Semitic phonology and morphology*. Ph. D. thesis, MIT. Published by Garland Press, New York, 1985.
McCarthy, J. J. (2007). Derivations and levels of representation. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology*, pp. 99–117. Cambridge: Cambridge University Press.
McCarthy, J. J. (2008). *Doing Optimality Theory*. Cambridge: Blackwell.
McCarthy, J. J. and A. Prince (1993b). Generalized alignment. *Morphology Yearbook*, 79–153.
Meringer, R. and K. Mayer (1895). *Versprechen und Verlesen*. Göschense Verlagsbuchhandlung.
Mielke, J. (2004). What ambivalent segments can tell us about the universality of distinctive features. Boston. LSA 78.
Myers, S. (1987). *Tone and the structure of words in Shona*. Ph. D. thesis, University of Massachusetts at Amherst, Amherst, MA.
Nespor, M. and I. Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.
Nijen Twilhaar, J. (1990). *Generatieve fonologie en de studie van Oostnederlandse dialecten*. Amsterdam: P.J. Meertens-Instituut.
Odden, D. (1980). Associative tone in shona. *Journal of Linguistic Research 1*, 37–51.
Odden, D. (1986). On the role of the obligatory contour principle in phonological theory. *Language 62*, 353—383.
Odden, D. (1987). Dissimilation as deletion in chuckchi. In A. Miller and J. Power (Eds.), *Proceedings from ESCOL*, Volume 3, Ohio, pp. 235–246. Ohio State University.
Odden, D. (2005). *Introducing phonology*. Cambridge: Cambridge University Press.

O'Roerke, E. (2008). Intonation in quechua: Questions and analysis. In *International Congress of Phonetics Sciences. Satellite workshop on 'Intonational Phonology: Understudied or Fieldwork Languages*.

Parker, S. (2002). *Quantifying the Sonority Hierarchy*. Ph. D. thesis, University of Massachusetts at Amherst.

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics 36*(1), 55–90.

Peng, L. (2013). *Analyzing Sound Patterns. An Introduction to Phonology*. Cambridge, UK: Cambridge University Press.

Pike, K. L. (1947). *Phonemics: A technique for reducing languages to writing*. Ann Arbor: University of Michigan Press.

Port, R. and A. Leary (2005). Against formal phonology. *Language*.

Prince, A. and P. Smolensky (1993). Optimality theory: Constraint interaction in generative grammar. Manuscript, Rutgers University and University of Colorado at Boulder. Available at ROA and published 2004 at Blackwell.

Radford, A. (2012). *Syntax. A Minimalist Introduction*. Cambridge, UK: Cambridge University Press.

Rehg, K. and D. Sohl (1981). *Ponapean Reference Grammar*. Honolulu: University of Hawaii Press.

Rhee, S. (2002). *Empty Nuclei in Korean*. Ph. D. thesis, University of Leiden.

Roettger, T. B., B. Winter, S. Grawunder, J. Kirby, and M. Grice (2014). Assessing incomplete neutralization of final devoicing in german. *Journal of Phonetics 43*(2), 11–25.

Sakel, J. and D. L. Everett (2012). *Linguistic Fieldwork. A Student Guide*. Cambridge, UK: Cambridge University Press.

Sapir, D. (1965). *A Grammar of Diola-Fogny*. Cambridge, UK: Cambridge University Press.

Schane, S. (1974). How abstract is abstract? In A. Bruck, R. A. Fox, and M. W. La Galy (Eds.), *Papers from the Parasession on Natural Phonology*, pp. 297–317. Chicago: Chicago Linguistic Society.

Sezer, E. (1986). An autosegmental analysis of compensatory lengthening in turkish. In L. Wetzels and E. Sezer (Eds.), *Studies in compensatory lengthening*, pp. 227–250. Dordrecht: Foris.

Sievers, E. (1881). *Grundzüge der Phonetik*. Leipzig: Breitkopf und Härtel.

Staroverov, P. (2014). *Splitting Theory and Consonant Epenthesis*. Ph. D. thesis, Rutgers University.

Steriade, D. (1988). Greek accent. a case for preserving structure. *Linguistic Inquiry 19*, 271–314.

Syed, N. (2012). Nasals and nasality in saraiki. *Language in India 12*(7), 323–347.

Topintzi, N. (2006). *Moraic Onsets*. Ph. D. thesis, University College London.

Trubetzkoy, N. S. (1939). *Grundzuge der Phonologie*. Prague: Travaux du cercle linguistique de Prague 7.

van Gelderen, E. (2017). *Syntax. An Introduction to Minimalism*. Amsterdam & Philadelphia: John Benjamins Publishing.

van Oostendorp, M. (1998). Bergen op zoom diminutives in phonological theory. In S. Barbiers, J. Rooryck, and J. van de Weijer (Eds.), *Small words in the big picture; Squibs for Hans Bennis*. Leiden: HIL.

van Oostendorp, M., C. J. Ewen, E. Hume, and K. Rice (Eds.) (2011). *The Blackwell Companion to Phonology*. Oxford: Blackwell.

Vennemann, T. (1988). *Preference laws for syllable structure and the explanation of sound change. With special reference to German, Germanic, Italian and Latin*. Berlin: Mouton De Gruyter.

Watson, J. (2011). Word stress in arabic. In M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice (Eds.), *The Blackwell Companion to Phonology*, Volume V, Chapter 124, pp. 2990–3018. Oxford, UK: Wiley-Blackwell.

Williams, E. (1976). Underlying tone in margi and igbo. *Linguistic Inquiry 7*, 463—484.

Woodrow, H. (1909). A quantitative study of rhythm. *Archives of Psychology 14*, 1–66.