

ENG 3360 - Introduction to Language Studies



# ENG 3360 - Introduction to Language Studies

*RÉGINE PELLICER*

UNIVERSITY OF TEXAS RIO GRANDE VALLEY  
EDINBURG, TX



ENG 3360 - *Introduction to Language Studies* by Régine Pellicer is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, except where otherwise noted.

# Contents

Introduction	1
Part I. Main Body	
Chapter 1: Language and linguistics	3
Chapter 2: Are humans unique?	37
Chapter 3: Language acquisition	38
Chapter 4: Phonology	64
Chapter 5: Morphology - The grammar of words	100
Chapter 6: Morphology- Word formation	139
Chapter 7: History of English	171
Chapter 8: Dialects	208
Technical notation and terms	209



Welcome,

Language studies cover a large variety of situations as language is embedded in every aspect of our lives. Finding a book that would study all the possible topics related to language is impossible. Therefore, with our UTRGV librarians, we have assembled a free book that covers the basic linguistics concepts you need to know for this course and other linguistics courses. Linguistics concepts such as phonology or language acquisition are not subject to last minute discoveries and the principles remain the same. We hope you enjoy learning more about linguistics thanks to this free resource.

Régine Pellicer





# Chapter 1: Language and linguistics

## Language and linguistics

### Chapter Preview

On language:

What do we mean by language?

What do we mean by grammar?

What competences do language users have?

What are some of the key features of language?

On linguistics:

What characterises linguistics as a discipline?

What are the key features of scientific investigation?

What criteria do we use to evaluate linguistic analyses?

How do we acquire data for linguistic analysis?

## 1.1 What do we mean by *language*?

Have you ever wondered whether language is a capacity unique to human beings? The answer obviously depends on what we mean by the word **language**. If you look in a dictionary, you will find at least two characterisations of language:

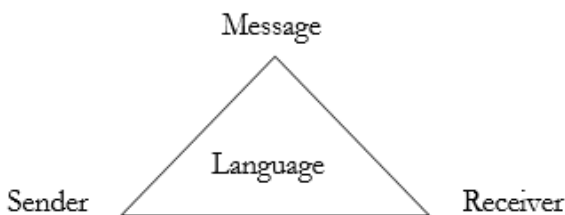
1. The system of human communication, using arbitrary signs (e.g. voice sounds, hand gestures or written symbols) in combination, according to established principles/rules.
2. Such a system as used by a nation, people or other distinct community/social group.

The main difference between the two meanings of the word *language* seems to be a difference between language in theory and language in use.

The first meaning focuses on language as a universal human phenomenon, language as a mental or cognitive phenomenon, which might be paraphrased by an expression like ‘language faculty’. In contrast, the second explanation focuses on language as a social phenomenon, i.e. the language faculty as it is expressed in individual languages like Mandarin, Portuguese, Malay or English. These two notions of language – as a mental phenomenon and as a social phenomenon – are captured, in English, in the one word, *language*. But in French, the language in which the father of modern linguistics, Swiss linguist Ferdinand de Saussure (1915/1974), conveyed his ideas about the nature of language, the two meanings of language are encapsulated in the words *langage*, referring to the language faculty, and *langue*, referring to particular languages. We detail the relationship between the two notions of language in the remaining sections of this chapter, as well as in the next chapter. But, for the moment, let us consider another way in which we might explain what is meant by the term *language*.

Typically, when we wish to explain what something is, we do so in terms of its **form** and/or **function**. For example, if you wanted to

explain to someone what a car is, you could describe it as a four-wheeled road vehicle powered by an internal combustion engine and designed to transport passengers. This explanation focuses on both form and function: it tells us what a car is made up of, and what it is for. Similarly, in explaining what language is, we can talk about its form and its function. In terms of form, language can be thought of as a **code**, i.e. a set of arbitrary signs (voice sounds, hand gestures, written symbols) and the rules for combining these into specific patterns, in order to express meaning. The crucial point to note here is the rule-governed (as opposed to random) nature of codes, and therefore of language. Codes comprise definite rules for combining signs into meaningful patterns, and this is why they allow meaning to be expressed through them. To communicate by means of a code, senders and receivers must share the same set of signs and rules for combining them. Likewise, language comprises a set of signs and the rules for combining them, in order to send and receive messages, as shown in Figure 1.1 below:



*Figure 1.1 Language as a code*

Given the analogy between language and codes, you might think that the primary purpose of language is to enable communication. After all, the purpose of codes is to enable communication between senders and receivers of coded messages. But, you would only be partially right. While codes and language are certainly both means of communication, the primary purpose of language, you

may be surprised to learn, is not communication but the expression of meaning, as highlighted in the quote below:

Language is a tool for expressing meaning. We think, we feel, we perceive – and we want to express our thoughts and feelings, our perceptions. Usually we want to express them because we want to share them with other people, but this is not always the case. We also need language to record our thoughts and to organise them. We write diaries, we write notes to ourselves, we make entries in our desk calendars, and so on. We also swear and exclaim – sometimes even when there is no one to hear us. The common denominator of all these different uses of language is not communication but meaning.

(Anna Wierzbicka, 1992, p. 3)

The question then becomes: how does language link meaning to expression? The answer is: through **grammar**, as depicted in Figure 1.2 below.

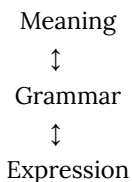


Figure 1.2. How language links meaning to expression

## 1.2 What do we mean by *grammar*?

Simply put, the grammar of a language comprises a set of signs (its sounds and words) and the principles, or **rules**, for combining these into meaningful utterances.

Given this concept of grammar, it follows that every human being who speaks a language knows its grammar, and it makes no sense for the speaker of any language to say *I don't know any grammar* or *My grammar is poor*. If you can function in a given language, then you know its grammar, i.e. the code for expressing meaning in that language.

One reason you may think you don't *know* the grammar of your language is that this knowledge is largely unconscious or implicit. An analogy might help to make things clearer. Think about your ability to walk or run. You know *how* to do it, but if someone asked you to explain what is involved in walking or running, you would probably end up with a severe case of paralysis by analysis, as would most people. Analysing what is involved in the process of walking or running is the job of the sports scientist, not the sportsperson. Similarly, all language users implicitly know the grammar of their language, and what the **linguist** seeks to do is to build a model of this mental grammar, i.e. of the implicit knowledge that speakers possess about the structure and use of their language.

At this point, it is worth correcting two common misconceptions about the nature of linguistics, which, we think, might be the source of misconceptions about the nature of grammar. The first is that linguistics is the study of language with the goal of learning to speak well. One of the most frequent questions linguists get asked, at parties and other social functions is: So, *how many languages do you speak?* or its variant *How many languages do you need to speak, if you want to become a linguist?* It certainly is true that some linguists are fluent in two or more languages. But the point is that a linguist is not someone who speaks several languages well. Such a person is a **polyglot** (from Greek *poly* meaning 'many' and *glotta* meaning

‘tongue’, or language). In other words, it is not necessary to speak several languages in order to be a linguist. Again, an analogy might help make things clearer. How many musical instruments do you need to play, if you wish to study music theory? The answer is none, since you can study music theory without knowing how to play any particular musical instrument. Obviously, when it comes to language, every linguist speaks at least one, by virtue of being human. But the point is that one does not need to speak any particular language, in order to do linguistics. (To find out how linguists manage to study the grammar of languages they do not speak, see sections 1.6 and 1.7.)

A second common misconception about linguistics is that it *prescribes* how people ought to speak, when, in fact, it *describes* how people actually speak. In other words, linguistics deals with what is, not what ought to be. This distinction between **prescription** and **description** is important in all disciplines, not just linguistics. Consider, for instance, the domain of ethics. If you wish to create a theory of ethics, you have two choices. You can either describe how people actually behave, in terms of the ethical choices they make. Or, you can prescribe how people ought to behave. Similarly, if you wish to create an economic theory, you can describe how people do actually behave, economically. Or, you can prescribe how you want them to behave. Most academic disciplines take the former stance, describing the world around us rather than prescribing behaviour. This is because prescriptivism tends to be based on the value system of some particular individual or group, and raises the question why that group’s value system should be the one guiding everyone else’s behaviour. This distinction between prescriptivism and description also applies to notions of grammar.

### Activity 1.1

Before you read the next section, write down what you understand by the term grammar. Then, compare your understanding of this word with that of your friends, your parents, your teachers. What similarities and differences can you see among the definitions you have collected?

One reason why the notion of grammar that we outlined at the beginning of this section may have surprised you is that many people think of grammar in prescriptive terms, i.e. as a set of rules that tell people how they should speak. Linguists refer to such grammars as **prescriptive grammars** in order to distinguish them from the kind of grammar that linguists are interested in. Prescriptive grammars are so named because they comprise a set of prescriptions about how to speak, which are based on nothing more than the idiosyncratic value judgments of the particular individual or group prescribing the rules. The rules below are all examples of prescriptive grammar rules:

Never start a sentence with *and* or *but*.

Never end a sentence with a preposition.

Never use a double negative.

Always say *I am not*, never *I ain't*.

If you ask the prescriber the reason for these rules, they are likely to answer that the best people (i.e. the people doing the prescribing or some group whom they wish to emulate) follow these rules, and therefore so should you. Not to speak in the fashion prescribed by them is considered bad form, and likely to be taken as a sign of the speaker's ignorance, stupidity, lack of class, etc. Prescriptive rules seek to control behaviour by telling people what to do/what not to do. Traffic rules are a good example of prescriptive rules. In contrast, descriptive rules state the patterns or regularities observed in the behaviour of entities of various sorts. The rules governing the motion of physical objects are a good example of

descriptive rules. Newton's Laws of Motion describe how objects in motion behave under different circumstances, not how they ought to behave. Similarly, the linguist's **descriptive grammar** is meant to model the principles, or rules, that speakers seem to be implicitly using, when they speak to one another.

### Activity 1.2

One of these statements is prescriptive, and the other is descriptive: which is which? How did you arrive at your answer?

1. Multilinguals should avoid mixing their languages when they speak.
2. Mixing languages is common in multilingual speech.

To summarise, then, prescriptive grammars tell people how they ought to speak, while descriptive grammars describe how they actually speak. In building descriptive grammars, linguists do not judge linguistic utterances to be correct/incorrect. Rather, they accept language users' intuitions about what sounds fine/odd in a given language variety, and seek to explain this asymmetry between what speakers consider acceptable/unacceptable, in the simplest and most general way possible (for more about the criteria used by linguists to evaluate alternative explanations, see section 1.6).



### 1.3 What competences do language users have?

In the preceding section, we emphasised that anyone who knows a language knows its grammar, i.e. the sounds and words of that language and the established conventions for combining these into meaningful utterances. In attempting to build a model of language users' mental grammar, linguists typically talk about two kinds of competence, or two kinds of knowledge that language users possess, namely, grammatical competence and communicative competence. **Grammatical competence** refers to language users' knowledge of the sounds and words of their language and how to combine these into **well-formed** utterances with the desired meaning. **Communicative competence** in turn refers to language users' knowledge of how to use language **appropriately** in order to carry out specific tasks and achieve the desired effects in specific communicative situations.

Here's an example to help you see the difference between the two kinds of competence. Consider the following utterance: *Can you tell me what time it is?* Speakers of English would judge this a well-formed expression, in contrast to the expressions *Can time it what tell you me is?* Or *You tell me can time is what it?*, which most speakers of English would judge as being ill-formed. This is grammatical competence at work. Now consider the following exchange between individuals A and B. A doesn't have a watch, and stops a passerby, B. A says to B, *I'm sorry I don't have my watch. Could you tell me what time it is, please?* B answers Yes, and walks off. Were you surprised by B's behaviour? If so, it's because your communicative competence tells you two things. First, that the question *Could you tell me what time it is, please?*, although structurally similar to questions like *Could you buy some milk on your way home?* or *Could you feed the dog tonight?*, which require just a yes or no answer, is different in function. Second, that in asking this question, A was not asking about B's willingness to tell

her the time, but was in fact requesting B to tell her the time. In other words, A wasn't saying *Would you be willing to tell me what time it is?* but *Please tell me the time.*

Since there may be a gap between what speakers say and what they actually intend, what is it that helps us reconstruct implicit meaning, and decide what forms are appropriate in a given communicative situation? The answer is: our communicative competence – our awareness of how **situational context**, including the setting, the purpose and the participants in a communicative exchange, affects the interpretation of any utterance. We discuss the importance of situational context in greater detail in Chapters 10 and 11, dealing with pragmatics, or meaning in action, and discourse, or language in use, respectively. For the moment, here's another simple activity to reinforce the idea that every language user, including yourself, possesses both grammatical and communicative competence.

### Activity 1.3

How would you interpret the following expressions?

NO BANANAS

NO BICYCLES

Both expressions have the same form. They both comprise two words: No followed by a plural noun – *bananas* and *bicycles*, respectively

In order to make sense of the meaning and the function of these expressions, we need to know the **context** in which they occur.

Since you haven't been told what this is, you could scan your memory for past encounters with these expressions. For example, you might have seen *No bicycles* on certain roads or paths, meant only for pedestrians or four-wheeled vehicles. This, combined with

you experience of other two-word expressions like *No dogs* or *No children*, might have led you to interpret the expression *No bicycles* as meaning ‘No bicycles allowed’. In short, this expression serves a directive function, prohibiting bicycles in the area where the sign appears.

The first expression (*No bananas*) may be a little harder to make sense of, but we can try to imagine a context in which you might encounter it. Suppose you saw these words on a handwritten sign at a fruit store. Would you infer the meaning ‘No bananas for sale’, on analogy with signs saying *No gas* at petrol stations that have run out of petrol to sell, during an oil embargo? If so, the function of this expression is declarative (telling us what is the case), rather than directive (telling us what to do), as in the case of the *No bicycles* sign.

#### Activity 1.4

On some buses and trains in South East Asia, the sign *No durians* can be seen. (Durians are a type of tropical fruit with a thick horny shell and a lingering smell, which some people like and others don’t.) What meaning and function would you infer from such a sign – the same as the *No bicycles* sign or the *No bananas* sign? Explain how you arrived at your answer.

What the activities above highlight are two important features of language use: the role of context in interpreting meaning, and the lack of one-to-one- correspondence between form and function – themes that we will be revisiting throughout this book.

## 1.4 What are some of the key features of language?

We now move on to consider five design features of language which, taken together, characterise language as a universally and uniquely human phenomenon, namely, arbitrariness, discreteness, compositionality, creativity, and rule-governedness.

### 1.4.1 Arbitrariness

We started this chapter with a dictionary definition of language as a system of human communication that makes use of arbitrary signs.

A sign is an indicator of something else, and signs can be divided into two kinds: arbitrary signs and non-arbitrary ones. Non-arbitrary signs have an inherent, usually causal, relationship to the things they indicate. The adage *where there is smoke, there is fire* is a fairly reliable observation, given that smoke-free fires tend to be a rarity. So we can say that smoke is a non-arbitrary sign of fire. Similarly, dark clouds are a non-arbitrary sign of impending rain, because there is a direct connection between the two.

Language is different from these phenomena in that the association between a linguistic form and the meaning it expresses is not inherent but established by convention. This explains why different languages have different signs representing a particular entity. Consider, for example, the animal that English speakers refer to as *elephant*. The linguistic sign for this entity is *gajah* in Malay. Why? Because these are the expressions that speakers of English and Malay have agreed (by established convention) should refer to this particular entity.

### Activity 1.5

Can you explain what Juliet means when she says to Romeo:

*A rose by any other name would smell as sweet.*

What Juliet seems to be highlighting is the arbitrary, or conventional, nature of linguistic signs. The flower that Juliet mentions happens to be called a *rose* because that is what speakers of English have agreed to call it, and not because there is any inherent connection between the word *rose* and the sweet-smelling entity that it designates. Malay speakers refer to the same entity as *bunga mawar*. In other words, Juliet is reminding us that the names that we give to things are not to be confused with the things themselves: linguistic signs should not be confused with the reality they represent.

There are, in fact, two levels of arbitrariness in language. The first is the arbitrary association between the **form** of a linguistic expression and its **meaning**. This level highlights the fact that there is no inherent link between the form of a word, i.e. the sounds that constitute a word, and the meaning of that word. For example, there is no inherent reason why the sequence of sounds [f], [r], [i], that gives us the word form *free*, should have the meanings that we associate with this word form, e.g. ‘not enslaved’ or ‘not busy’ or ‘without payment’. The same meanings could just as easily be expressed by some other sound sequence, e.g. [gip] or [gub] or [bub]. It just so happens that speakers of English have agreed to

associate this particular sound- sequence [fri] with this particular set of meanings.

The second level is the arbitrary association between a **linguistic sign** and its **referent**. This level highlights the fact that there is no inherent link between a linguistic expression and the entity that it designates. For example, there is no inherent reason why the word *tree* should refer to the entity that it in fact refers to. There is no logical reason why the word *tree* could not refer to Juliet's rose, except that speakers of English have agreed by convention to use the words *rose* and *tree* in the ways that they do.

This double **arbitrariness** between linguistic forms and their meanings, and linguistic signs and their referents is one of the characteristics of language, first discussed by Saussure (1915/1974) in his *Cours de linguistique générale* (*Course in general linguistics*).

## 1.4.2 Discreteness and compositionality

Discreteness and compositionality are two sides of the same coin. **Discreteness** has to do with the fact that speakers of all languages can identify distinct elements in their languages, such as the different words in a sentence, and the particular sounds in a word.

Consider the utterance *The cats at on the mat*. As a speaker of English, you have no difficulty identifying distinct elements in the sound stream making up this utterance – you hear both the individual sounds in every word and the individual words making up this utterance. This ability to distinguish discrete units in a stream of language is in fact one of the things you become aware of, as you acquire a language. You realise that you can make out individual sounds and words in the stream of language you are hearing.

While discreteness refers to our ability to perceive distinct units within larger stretches of language, **compositionality** refers to our perception that larger units of language are composed of smaller units. Speakers of English perceive the utterance

*The cats at on the mat* as being composed of the words *The*, *cat*, *sat*, *on*, *the*, and *mat*. Similarly, they perceive each word as being composed of individual sounds, e.g. the word *cat* as being composed of three sounds, which we can represent by means of the symbols [k], [æ] and [t].

Putting discreteness and compositionality together, we can say that speakers perceive language as being composed of discrete units which stand in a part-whole relationship to one another. As you continue to work your way through this book, you will learn more about some of these units and their relationship to one another.

### 1.4.3 Creativity

Like arbitrariness, which functions on two levels, **creativity** in language manifests itself in two ways. First, language itself is creative in that it enables the expression of new meanings, as and when they are needed, for example through the creation of new words. As will become clearer in the next chapter, language is an organic entity, which changes, grows and dies, just like the human beings that use it. This intrinsic creativity is one of the factors underlying the phenomenon of **language variation**, across time, space and situations of use, which we also explore in greater detail in the next chapter.

What is interesting about the creativity of language is that it involves the “infinite use of finite means,” as highlighted by the 19th century German philosopher Wilhelm von Humboldt (1836/1999: 221). All languages have a finite repertoire of sounds. Yet, from this finite set of sounds, every language is able to build a vast range of words, to serve the needs of its speakers. From this vast vocabulary or **lexicon** (the inventory of words in a language), in turn, an even greater number of sentences can be constructed to express every possible meaning which speakers wish to convey.

The infinite use of finite means also characterises the second way in which language is creative, namely, speaker's use of language. Think about your own language use. When you wish to express some meaning, do you have a limited repertoire of sentences stored in memory for immediate recall? Or do you find yourself uttering novel sentences, as and when you need them? Consider, for example, the sentence below:

*The little old lady who tried  
to carry the Golden  
Retriever disguised as her  
son into the 601 bus was  
told off by the commuter  
holding a fainting bald  
eagle by its left foot.*

We are pretty sure that you will not have encountered this sentence before, since we just made it up ourselves. In other words, this sentence is novel both for you and us. Yet, neither of us has any trouble making sense of the sentence. What this suggests is that speakers of all languages can both understand and produce utterances that they have never before heard or spoken. The reason we are able to do this is that all languages have a limited repertoire of sounds and words, which can be combined to form an infinite number of meaningful utterances through a limited set of rules. This brings us to the fifth characteristic of language that makes it universally and uniquely human – its rule-governed nature.

#### 1.4.4 Rule-governedness

We started this chapter by comparing language to a code. We said that codes allow the expression of meaning because they comprise a set of signs as well as rules for combining the signs into meaningful patterns. Paradoxically, it is the rule-governed nature of language that makes language use a creative enterprise. It's easy to mistake



creativity for a case of anything goes, for absence of rules. That, however, would only result in total confusion, or in the case of language, random noise. What language use involves, as any other creative endeavour, is the combination of a finite set of resources in new and different ways, in order to create new meanings.

When linguists explore the rules that govern language, they tend to focus on two kinds of organisation: **linear order**, or sequencing, and **hierarchical order**, or part-whole relationships. We talked briefly about hierarchical order when discussing compositionality as a characteristic of language, in section 1.4.2. Hierarchical order has to do with constraints in the part-whole relationships assumed by discrete units of language. Linear order in turn has to do with constraints in the sequencing of language elements. Here's an example of linear order in terms of the sound-sequences that form words. Consider the three sounds in the English word *cat*, represented by the symbols [k] , [æ] and [t]. Each of these sounds has no meaning in and of itself. But each can be combined with the other two to form meaningful units at a higher level, i.e. words. What's interesting here is that mathematically speaking, three sounds can be ordered sequentially in six ways, as shown below (in linguistics, an asterisk preceding an example indicates a non-occurring form in a particular language):

[kæt] [tæk] \*[ktæ] \*[tkæ] \*[ætk] [ækt]

Yet, of these six sequences, only three are sanctioned by the sound- sequencing rules of English: [kæt] *cat*, [tæk] *tack*, and [ækt] *act*. These data reflect the intuitions of speakers of English about the possible sound- sequences of English. That is to say, if you asked speakers of English which of the six forms above represent English-sounding words, as opposed to foreign-sounding ones, their judgments would match our observation that that [kæt], [tæk] and [ækt] are fine, but not [ktæ], [tkæ] and [ætk]. In other words, the grammar of English disallows sequences like \*[ktæ], \*[tkæ] and \*[ætk]. Once you have read Chapter 6, dealing with the sound systems of languages, you should be able to revisit these data, and

articulate a rule that explains why words like \*[ktæ], \*[tkæ], and \*[ætɕ] are non-occurring in English.

In discussing these and other examples about the words of any language, it's important to keep in mind the difference between possible and actual words. For any language, there are words which could exist in it, but for some reason don't. **Possible words** are words that could be part of a particular language, i.e. that are sanctioned by the grammar of the language, but just happen not to. This means that there are accidental lexical gaps (word gaps) in all languages. For instance, if you asked English speakers whether the sound sequence [fik] is a possible English word, they would answer yes. In contrast, if you asked the same set of speakers whether [bnik] is a possible English word, they would probably answer no. What [fik] and [bnik] share is that neither is an **actual word** of English. The difference between them is that [fik] is a possible word of English, whereas [bnik] is not. The reason is that the grammar of English allows the sound-sequence [fik] but not the sound-sequence [bnik]. Once again, you should be able to revisit these data to articulate a rule that explains why sound-sequences like [bnik] [bnæk] and [bnag] are not possible words of English.

Linear order is important for studying not just possible sound-sequences in words, but also possible word sequences in sentences. Consider, for example, the five-word sentence *The cat licked the boy*. Mathematically speaking, there are 120 ways to sequence five words. Most of these sequences, however, result in nonsensical strings like *\*Boy the licked cat the* or *\*Licked cat boy the the*. One results in a different sentence, *The boy licked the cat*. We will be looking more closely at the rules governing word order in sentences. Meanwhile, the sound-sequence and word-order examples above illustrate the rule-governed nature of language. All languages have rules that enable them to convey meaning. The **grammar** of any language is a statement of the coding rules of that language, shared by the speakers of the language. It is these rules that linguists seek to capture and model in descriptive grammars. How they go about doing this is the subject of the remaining sections of this chapter.

## 1.5 Linguistics: the science of language

In section 1.2, we corrected two common misconceptions about linguistics, pointing out that linguistics is the *scientific* study of language as a human phenomenon. In this section and the next, we explain what we mean by science, and highlight the features that characterise linguistics as the science of language.

Simply put, **science** is the art of looking for patterns, and for ways of explaining them. Scientists look for regularities in the world around us, and propose analyses for these regularities in the simplest, most general and most objective way possible. The characterisation of science as an art is deliberate, on our part, given the tendency to treat science and art as mutually exclusive domains. Certainly, in everyday usage, the word *science* is often used as a cover term for domains of study such as physics, chemistry and biology to distinguish it from areas like history, geography or literature. Science, however, refers not to an area of study but to a way of studying. In other words, any phenomenon can be the object of scientific study.

Every one of us is a scientist to the extent that we look for regularities or patterns in the world around us, and seek to explain them in terms of theories that we hold, either implicitly or explicitly. For example, if you've ever been puzzled by a friend's behaviour, it's because you have a theory about how friends in general behave, and how your friend in particular behaves, within which the quirky behaviour that you observed does not seem to fit. If you then decided to investigate the reason for your friend's odd behaviour, in order to better understand it, you might have ended up having to amend your theory about your friend's behaviour, in particular, and/or friends' behaviour, in general. Observing and analysing situations, critically evaluating alternative analyses, and amending (or rejecting) an analysis on the basis of new (or conflicting) data are elements central to scientific investigation, the topic of the next section.

## 1.6 The nature of scientific investigation

Any investigation, including scientific investigation, can be thought of in terms of three parameters: the object of investigation, the method(s) of investigation, and the purpose of investigation. These three features correspond to three questions:

- What do you wish to study?
- How will you study it?
- Why do you wish to study it?

Being the science of language, linguistics has an object of investigation, namely, language as a universal human phenomenon. Its method is empirical. That is, the conclusions drawn in linguistics, as in any scientific investigation, are based on observation and experience rather than on intuition or pure reasoning. The purpose of linguistic analysis is to explain the nature of language, in order to be able to answer questions like What constitutes knowledge of language? How is language acquired? and How is it used? In the next three subsections, we look at the object, method, and purpose of scientific investigation a little more closely.

### 1.6.1 The object of investigation

Typically, scientific investigation deals with observable phenomena. In other words, science is interested in factual claims, in assertions which can be demonstrated to be true or false, based on empirical evidence. The difference between fact (something that you know to be true) and opinion is that the latter cannot be demonstrated to be true or false because it has to do with matters of judgment and taste. Compare, for example, the following two statements:

*All Singaporeans speak English.*

*All Singaporeans speak good English.*

The first statement is a factual claim. Its truth can be verified by checking whether or not all Singaporeans do in fact speak English. If you find at least one Singaporean who does not speak English, you will have demonstrated that this statement is false and therefore needs to be amended to a statement like Not all Singaporeans speak English or Most Singaporeans speak English. The second statement is identical to the first, except for the presence of the word good, which expresses a value judgment. In the absence of a shared theory about what makes a particular variety of English good or bad, what we have here is an expression of someone's opinion or taste, which does not fall within the realm of science, including the science of language.

## 1.6.2 The method(s) of investigation

Scientific investigation typically involves three key activities: observation, analysis and argumentation. Observation consists in looking for patterns in phenomena of various kinds. These patterns are then accounted for in terms of analyses, which typically comprise a set of theoretical constructs, a set of rules describing the observed patterns, and a set of representations that model the observed patterns.

It's important to bear in mind that a theoretical construct is a hypothetical entity assumed by the scientist in order to explain an observed pattern. Gravity, for example, is one of the best known and most useful theoretical constructs, given its immense explanatory power. Newton's theory of gravity explains observations both on earth (why things fall) and in space (why planets don't collide into one another). The constructs that linguists use to explain linguistic phenomena include assumed entities like nouns, verbs, phrases, morphemes and phonemes. We will meet these and other constructs in the remaining chapters of this book, as we explore the structure and use of language in greater detail.

## *Criteria for evaluating scientific and linguistic analysis*

Meanwhile, having looked at the first two activities entailed by the method of scientific investigation (observation and analysis), let's consider the third, namely, argumentation. **Argumentation** is necessary in order to critically evaluate alternative analyses. As you know from your own observation and analysis of the world around you, there can be multiple perspectives and explanations for the same phenomenon. How do we decide which analysis is best? Clearly, any disciplinary community needs shared criteria of evaluation. Six criteria that are commonly used to evaluate scientific and linguistic analyses are purpose, accuracy, simplicity, generality, objectivity, and internal consistency. We now deal with each one in turn.

**Purpose.** Any analysis needs to be motivated, that is to say, it must serve a purpose. The fundamental purpose of scientific analyses is the explanation of observed phenomena, which allows for the prediction of these and other phenomena. A scientific analysis that doesn't explain anything isn't a very useful analysis. Neither is an analysis that explains some trivial fact. The latter raises the "So what?" question: So, *you discovered X*. So *what*? In other words, scientific researchers need to show how their work contributes to the extant knowledge pool, either by providing new insight on an old problem/puzzle, or by highlighting new frontiers worthy of investigation. If the analysis merely rehashes what is already known, the reaction is likely to be, *Tell us something we don't already know*.

**Accuracy.** One of the key reasons that we seek explanations is in order to make predictions. Obviously, we want our predictions to be correct rather than incorrect. So, one of the basic requirements of an analysis is that it be consistent with our observations, i.e. that it not yield incorrect predictions. For example, if an analysis predicts that X will occur, but what we observe is not X but Y, we have an erroneous analysis, which clearly has to be rejected or amended to accurately capture Y. Analyses typically comprise a set of constructs and a set of rules, expressed as hypotheses

to be verified. A **hypothesis** is a prediction based on an initial set of observations. The purpose of a hypothesis is to be verified (i.e. either confirmed or disconfirmed) through empirical testing. If there is a mismatch between what the hypothesis predicts and what we in fact observe, the hypothesis is disconfirmed, and has to be rejected. In contrast, if what we see is in fact what the hypothesis predicts, we accept the hypothesis provisionally, putting it through further rounds of testing against new sets of data/observations. In both cases, hypotheses help us progress with our investigation. A disconfirmed hypothesis lets us know that the path we took was wrong, and that we therefore should take an alternative route. A confirmed hypothesis becomes a reliable piece of knowledge upon which we can build other hypotheses. .

**Simplicity.** You may have heard of **Occam's Razor** (also known as the law of parsimony or the law of economy), a principle attributed to 14th century logician William of Ockham. Simply put, Occam's Razor states that the simplest explanation is the best. In other words, when multiple explanations are available for the same phenomenon, the least complicated version is to be preferred, i.e. the one that makes the fewest possible assumptions. For example, hoof beats can be a sign of approaching horses or zebras. According to Occam's Razor, if one heard hoof beats on a farm, horses rather than zebras would be the preferred explanation, since it requires the fewest assumptions, given what we know about horses and zebras relative to farms. Similarly, burning bushes could be caused by a discarded cigarette or a landing alien spacecraft. Which explanation would you choose, using the simplicity criterion?

**Generality.** The generality criterion requires an analysis to cover the widest possible range of observations. Why? Because the more general a theory or analysis is, and so the more observations it covers, the fewer theories we need. An analogy might help. One of the reasons we like credit cards is that they can be used for a whole host of transactions, both online and offline. If credit cards didn't have such general coverage, we would need a whole variety of financial instruments to take care of the different transactions we

wish to perform. One of the best known theories, Newton's theory of gravitational forces, fulfils the generality criterion by explaining observations both here on earth and in space. A less general theory would account only for some observations (e.g. the ones here on earth) but not others. Similarly, if Newton's theory could only account for some falling bodies (e.g. apples) but not others (e.g. cannonballs or human beings), it would have to be rejected for not being general enough, i.e. for being unable to explain all our observations relating to falling objects.

**Objectivity.** Scientific knowledge is objective in the sense that it is supported by observation and experience rather than by intuition or pure reasoning. Maximal objectivity refers to the use of different types of evidence, or independent sources of evidence, to support a conclusion. This use of multiple types/sources of evidence to establish a conclusion is in turn known as **independent corroboration**. When various strands of evidence all point towards the same conclusion, we have **converging evidence** for that conclusion.

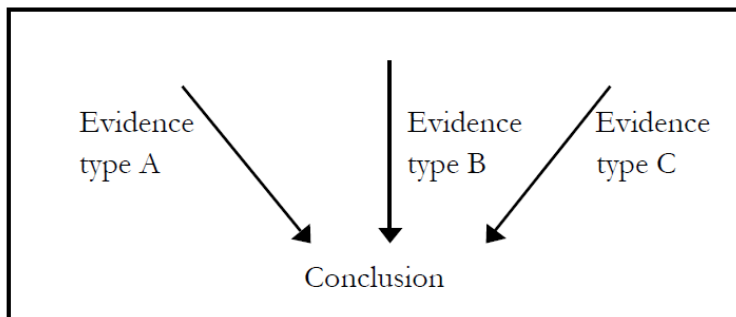


Figure 1.3. *The method of independent corroboration*

Clearly, the more independent reasons there are for accepting a conclusion, the more reliable that conclusion is likely to be. Why? Because each type of evidence would have to be countered for the conclusion to be refuted. Think of two cables: one comprises



just one strand of wire, the other comprising multiple strands of wire. Which cable is likely to be stronger? The latter, because one would have to cut through several strands in order to sever the cable. Here's another example. Let's say you suspect Jim Jones of murdering his wife. All you have is a gut-feeling that Jones is guilty. No court would accept this subjective evidence as sufficient reason to convict Jones of murder. Now, let's suppose you manage to obtain eyewitness testimony, placing Jones at the scene of the crime. This would constitute objective evidence. But this evidence could be weakened, if one or more of the eyewitnesses were shown to be unreliable. One way of strengthening your case, then, would be to find an independent source of evidence, implicating Jones in the murder of his wife, e.g. forensic evidence, so that even if the eyewitness testimony were impeached, you would have another source of evidence leading to the same conclusion.

**Internal consistency.** Whereas the accuracy criterion requires analyses to be free of error, and thus not to yield incorrect predictions, the internal consistency criterion requires analyses to be free of logical contradiction. An analysis that states/predicts both X and not-X would be internally inconsistent, or self-contradictory, because X and not-X are mutually exclusive terms. For example, an entity can be short or tall, but not both, at the same time. It may seem hard to believe that a careful researcher could construct an analysis in which one part of the analysis contradicts another. But the possibility remains, especially with longer analyses.

### *Linguistics data and ways of acquiring them*

Having looked briefly at analysis and argumentation, we return to the central activity in the method of scientific investigation, observation. What is it that linguists observe, and how do they go about collecting their data?

To start with the first question, linguistics data comprise language behaviour as well as speakers' intuitions about language, produced

either by themselves or others. These data can be obtained through observation of naturally-occurring situations or manipulated situations, as in experimentation. In the case of spoken language, naturally-occurring data are typically obtained by observing spontaneous communicative events like conversations, paying attention to the language behaviour of each participant. For example, who opens/closes the conversation? How do speakers signal their wish to change the topic? Who interrupts whom, when and how? Given the ephemeral nature of spoken language, researchers these days often use video or audio technology to record such exchanges. They can then watch/listen to them, transcribe them, and analyse them more accurately. In this sense, analysing written language is a little easier, given that writing leaves a more permanent record than speech. Since the 1980s, however, linguists at various research sites have been building databases or corpora of both spoken and written language (e.g. The British National Corpus, the International Corpus of English) that linguists wishing to investigate naturally-occurring language can access.

In contrast to observing naturally-occurring linguistic behaviour, observation of elicited data typically involves some amount of manipulation, so as to obtain the language behaviour that the researcher wishes to investigate. For example, back in 1958, psycholinguist Jean Berko Gleason wished to investigate the acquisition of the plural rule among English-speaking toddlers, pre-schoolers, and first graders. Gleason (1958) designed an experiment, known as the Wug test, which called for, among other things, the use of the plural. She and her fellow researchers showed each child in the experiment a picture of a pretend creature, and told the child, This is a wug. Next, the child was shown a picture with two of these pretend creatures, and told, Now there are two of them. There are two.... If you'd like to learn more about Gleason's experiment and her findings, turn to the references at the end of this chapter.

Here's one more example of observation of elicited data, this time involving adult speakers of English in New York City, in the

late sixties. You may have noticed that in words like four, fourth, card, and car, some speakers of English pronounce the /r/ sound represented by the letter 'r' in these words, while others drop it. While he was still a graduate student, sociolinguist William Labov (1966) hypothesised that the pronunciation of /r/ in New York English was not random, but correlated to social-class. He predicted that people of higher social status would pronounce /r/ more frequently than people of lower social status. To test his hypothesis, Labov investigated the speech of employees in three Manhattan department stores – an expensive, upper-middle-class store; a mid-priced, middle-class store; and a discount working-class store. In each store, Labov asked various employees for the location of a product which he knew was available on the fourth floor of the store. His question elicited the utterance fourth floor from the employees, and he was thus able to compare their pronunciation. For more about Labov's experiment and findings, turn to the references at the end of this chapter.

To recap, then, observation of language behaviour, whether naturally- occurring or elicited, is an integral part of the method of linguistic investigation. But language behaviour is only one kind of data that linguists are interested in. Other data include speakers' intuitions about their language, typically obtained through questionnaires or interviews, or through **introspection**. The latter refers to linguists consulting their own intuitions about a language, when they happen to speak the language they wish to investigate. As mentioned in section 1.2, however, linguists do not need to speak the language they wish to investigate. If you're wondering how linguists can be credited with providing trustworthy descriptions of languages that are alien to them, think about how doctors discover what is ailing their patients. Doctors don't need to be suffering from a particular disease in order to investigate it. What they do is examine patients and ask questions that allow them to infer the nature and cause of the disease. Likewise, linguists examine available language behaviour and ask **native speakers** about their intuitions concerning language. This is what we (the authors of

this book) would have to do, if we wished to propose a credible description of Singapore English or Jamaican English, for example, not being native speakers of these varieties of English ourselves. We would both observe the language behaviour of native speakers of these varieties and obtain their intuitions about their language use in order to propose an analysis of the grammar of Singapore English and of Jamaican English, respectively. Our own use of English would tell us only about the varieties of English that we speak, and little or nothing about the variety of English spoken by Singaporeans and Jamaicans.

### 1.6.3 The purpose of investigation

As mentioned in section 1.5, the purpose of scientific investigation is to explain observed phenomena as simply, generally, and objectively as possible, so that we can not only satisfy our curiosity about the world around us, but also adapt to it by being able to predict events reliably.

At the beginning of this chapter, we also highlighted that in defining/exploring any object, it is possible to focus on form or function. Needless to say, if we are to achieve a holistic understanding of the nature of language, we need to consider both aspects, since function influences form, and vice versa. Given the complex interplay between form and function, it would be virtually impossible to find a branch of linguistic investigation that adopts an exclusively formal or functional perspective.

Moreover, given how language permeates human activity, its study affects and is in turn affected by virtually every other area of human interest. The study of language today tends to be approached from the perspective of many different sciences and professional fields, from which linguists draw knowledge and to which they contribute practical and theoretical insight. Some of these interdisciplinary areas of research include:

- Anthropology: language and culture; ethno-linguistics
- Child studies: language acquisition and development
- Computer science: programming languages; machine translation; “thinking” machines; computational linguistics
- Education: language teaching
- Dentistry: orthodontics and speech
- History: language change over time; historical linguistics
- Law: forensic linguistics
- Mathematics: formal systems of communication
- Medicine: language, mind and brain; neurolinguistics; language pathology
- Music: the singing voice
- Philosophy: language and reasoning
- Physics: voice acoustics; speech synthesis; voice recognition
- Psychology: language and cognition; psycholinguistics
- Rhetoric: language and persuasion
- Sociology: language, ideology, and power; sociolinguistics; critical discourse analysis

In this chapter, we looked briefly at the nature of **linguistics** as the science of language, to systematically observe and explain language structure and use. We also considered two notions of language – as language faculty (Saussure’s *langage*) and its manifestation in individual languages (Saussure’s *langues*). We might think of these two notions of language as two sides of the same coin. That is to say, no matter what the differences between individual languages, there must be certain shared universal features, given that every language must conform to the constraints of the human brain and of human communicative needs. Conversely, the study of language as a universal human phenomenon can only proceed through the study of individual **languages**. The next chapter considers this relationship between language and languages more closely.

# Food for thought

## *On language*

“It is a very remarkable fact that there are none so depraved and stupid, without even excepting idiots, that they cannot arrange different words together, forming of them a statement by which they make known their thoughts; while, on the other hand, there is no other animal, however perfect and fortunately circumstanced it may be, which can do the same.”

René Descartes (1637).

*Discourse on the method.*

“If a lion could talk, we could not understand him.”

Ludwig Wittgenstein (1921).

*Tractatus logico-philosophicus.*

“It is certainly the business of a grammarian to find out, and not to make, the laws of a language.”

John Fell (1784). *An essay towards an English grammar.*

Reprinted 1974, London: Scolar Press.

## *On the nature of scientific investigation*

“How do we know?”

“The act of observing affects the observed.”

##

Werner Heisenberg

“It’s a commonly held view that ‘facts’

are just lying about in the world, and the way we make theories is by collecting these facts and then seeing what theories they lead to. Nothing could be further from the truth: in a way, there are no facts without theories. One might even define a theory as – in part – a framework that tells you what a fact is.”

R

oger Lass  
(1998).  
Phonology.  
An  
introduction  
to basic  
concepts.  
Cambridge:  
Cambridge  
University  
Press, p. 6.

“... most writing about [science] focuses only on the *answers*. People cannot make sense of answers if they do not first understand the *questions*. Solutions only have meaning if one has a firm grasp of the problems being addressed, and of why these problems matter.

Why, for instance, does it matter if the earth revolves around the sun or the sun around the earth? In most physics books, and in most classrooms, this is presented as a problem in celestial geometry: Is it the blue dot or the yellow one at the center? With virtually no sense of context we are told that Copernicus finally

“solved” this problem by placing the yellow dot in the central position. To most students the whole exercise appears little more than an abstract mathematical game.

Yet the issue matters greatly. The question of whether the sun or the earth is at the center of the cosmic system is not just a matter of celestial geometry (though it is that as well), it is a profound question about human *culture*. The choice between the geocentric cosmology of the Middle Ages and the heliocentric cosmology of the late seventeenth century was a choice between two fundamentally different perceptions of mankind's place in the universal scheme. Were we to see ourselves at the center of an angel-filled cosmos with everything connected to God, or were we to see ourselves as the inhabitants of a large rock purposelessly revolving in a vast Euclidian void? The shift from geocentrism to heliocentrism was not simply a triumph of empirical astronomy, but a turning point in Western cultural history.”

Margaret

Wertheim (1997).

*Pythagoras' trousers. God, physics, and the gender wars.* London: Fourth Estate, p. xii.

“... we [the non-scientists] may not be able to follow the details of the scientists’



proofs, but we are entitled to explanations we can understand.”

Lisa Jardine (2000). *Ingenious pursuits: building the scientific revolution*.

London: Abacus, pp. 364-365.

“... to my mind by far the greatest danger in scholarship (and perhaps especially in linguistics) is not that the individual may fail to master the thought of a school but that a school may succeed in mastering the thought of the individual.”

Geoffrey Sampson (1980). *Schools of Linguistics. Competition and evolution*.

London: Hutchinson, p. 10.

## Further reading

Brinton, Laurel J. (2000). Chapter 1. The nature of language and linguistics. In *The structure of modern English: A linguistic introduction*. Amsterdam: John Benjamins, pp. 3-11. Crystal, David (1986). Chapter 2. What linguistics is. In *What is linguistics?* (4th ed.) London: Edward Arnold, pp. 24-52. In *Science We Trust*. (2002, December). *Scientific American*. Editorial, p. 4. Napoli, Donna Jo (2003). Part I. Language: The human ability. In *Language matters: A guide to everyday thinking about language*. New York: Oxford University Press.

## References

Gleason, Jean B. (1958). The child's learning of English morphology. *Word* 14, 150-177.

Humboldt, Wilhelm von (1836/1999). *On language. The diversity of human language-structure and its influence on the mental development of mankind*. Cambridge: Cambridge University Press.

Labov, William (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Saussure, Ferdinand de (1915/1974). *Course in general linguistics*. Eds. Charles Bally and Albert Sechehaye in collaboration with Albert Reidlinger. Translated from the French by Wade Baskin. London: P. Owen.

Wierzbicka, Anna (1992). *Semantics, culture, and cognition: Universal human concepts in culture-specific configurations*. Oxford/New York: Oxford University Press.

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.

# Chapter 2: Are humans unique?

Use the link here to access *Are Humans Unique?*, chapter 1, via the UTRGV Library.

<https://www-taylorfrancis-com.ezhost.utrgv.edu/chapters/mono/10.4324/9781315880358-11/humans-unique-martin-atkinson-iggy-roca-david-kilby?context=ubx>

# Chapter 3: Language acquisition

## Language and speakers

### Chapter Preview

Who is a native speaker?

Who is a multilingual?

Are there universal stages of language development?

What are language loss and language death, and why do they happen?

### 3.1 Introduction

We started this book attempting to define what we mean by language. We highlighted two notions of language, language as a universal human faculty (captured in the French word *langage*) and language as a social phenomenon comprising the range of languages spoken by human beings around the globe (corresponding to the

French word *langue*). Having discussed issues pertaining to different languages, language structure, meaning and use, we conclude our exploration of the nature of language by considering the users of language, and the seemingly trivial issue of what to call them. In the process, we address the issue of linguistic **taboo** in an area where it might be least expected: the language of language itself.

## 3.2 The natives

To be considered a *native* of a country, all you need is to be born in that country. The question here is: what does it take to be considered a **native speaker** of a language? The problem with the definition of the compound word *native speaker* lies in its modifier: how exactly does the stem *native* modify the head *speaker*? Judging by the flurry of literature addressing the definition of *native speaker*, there is no simple answer to this puzzle.

Take one example. Due to perceived racist connotations of the term *Indian*, North-American Indians are currently called *Native Americans*, a label that appears to suggest that people of non-Indian ethnicity who are born in the United States are not native Americans. In its current use, upper-cased *Native American* is in fact a hyponym of the superordinate term *American*, which includes both native-born Americans (only some of whom are Native Americans) as well as naturalised citizens.

In some cases, the definition of native speaker appears straightforward: a Briton who is born and bred in Britain, and is a monolingual speaker of English, is a native speaker of English. But what are we to make of the following situation? Born in France to monolingual French parents, Mathilde lived in France until the age of seven, then settled with her parents in a monolingual English-speaking country, where she attends school in English and has no contact with French except through her parents. Based on her interaction with her peers and teachers at school and in the

playground, Mathilde acquires a mastery of English that is indistinguishable from that of her “native” schoolmates. Her French in turn is restricted to interaction with her parents.

Is Mathilde still a native speaker of French, even though her command of the language may not be native-like? Is Mathilde now a native speaker of English, since her command of the language is native-like? The answer to these questions holds a clue to the definition of native speaker. This can be summed up in the adage, once a native speaker, always a native speaker. In other words, being a native speaker has more to do with birth-right than linguistic proficiency. You are either a native speaker or you are not. You can neither become a native speaker, nor stop being one, as evidenced by the strangeness of formulations like I became a native speaker of English at the age of seven or I stopped being a native speaker of French in my teens.

To return to Mathilde’s situation, we could describe her as a native speaker of French and a multilingual speaker of English. We use the term **multilingual** to designate users of more than one language, thus including **bilinguals**, trilinguals, and so on. But this label does not entirely capture her native(-like) command of English. This is especially so given the fact that the labels bilingual/multilingual are often used synonymously with semilingual, as we shall see in section 12.3 below.

### 3.2.1 Language acquisition

All children acquire the language(s) that are spoken in their environment, and all children acquire language in the same way and at the same pace. At all stages of typical language development, universal patterns can be found. For example, all children start by producing sequences like [gugugu], which give its name to the so-called **cooing stage**. Sequences like [b a b a b a], [d a d a d a], or [dididi] follow, in the **babbling stage**, but not sequences like

\*[fæfæfæ]. Child preference for sequences like [b a b a] and [d a d a] is what explains the prevalence, in many different languages, of words constituted by a reduplicated sequence of [+stop +labial] or [+stop +coronal], followed by an open vowel, to designate mummy and daddy. Since time immemorial, parents all over the world have been eager to assign meaning to their children's productions, and preferably meanings that involve themselves as referents.

All children's babbling reflects uses of **pitch**, as well as other core components of any human utterance, in sequences of rises vs. falls, stressed vs. unstressed syllables or high-pitched vs. low-pitched syllables. These essential components of language are in fact the first ones used by children to communicate meanings, such as feelings, demands or queries, in the absence of words.

After the babbling stage comes the **one-word stage**, where children's utterances consist of single words only, all of which are lexical words. Common one-word utterances among English-speaking children include Doggy, Ball, Drink – closely followed, of course, by No!

### Activity 3.1

In the speech of some English-speaking children, the following pronunciations are found: sat [tæt] team [tim] tar[tʰ] Sue [tu] Sam [tæm] see[ti] Do these pronunciations match what you now know about child speech? Why?

The **two-word stage** then follows, signalling the beginning of syntax. Collocations in child speech are as significant as in adult

speech: child utterances like *Dolly give* and *Give dolly* mean different things.

As their linguistic development continues, all children go through stages where they apparently make mistakes like saying *drinked* and *comed* for *drank* and *came*. In fact, such mistakes signal the emergence of morphological **rules** in child speech. That is to say, mistakes such as the ones above suggest that the child has acquired the rule for regular past tense formation in English, but overgeneralises it to irregular verbs. **Overgeneralisation**, or *overextension*, is apparent in other areas like word meanings, where a word like *moon* can be used to designate a full moon, a waxing crescent, a banana or a lemon wedge, on the basis of perceived similarities in form shared by all these referents.

Instances of overgeneralisation in child speech in fact constitute solid evidence against the popular view that children learn language through simple imitation of adult speech: the overgeneralised child forms do not occur in adult speech and cannot therefore be imitated. Rather, what children appear to do is to filter the speech they hear around them according to patterns that they progressively uncover. Child strategies to acquire command over the system behind adult uses of language are in this sense no different from those used by a code-breaker assigned the task of cracking a code. The difference between the two tasks is that children don't need to figure it out all by themselves. Adult and other older language users guide the child by means of **motherese**, the language that nurtures the development of language. Motherese (also known as **child-directed speech**, a euphemism that avoids the female denotation of the original word) mirrors the linguistic abilities that are perceived in the child, and progressively expands these. At the two-word stage, one example of an exchange involving motherese is shown below. The mother, who is trying to get the child to nap, pops a toy dog snugly into the child's bed and pats it:

Mother. *Shhh, the doggy is asleep!*

Child. *Doggy sleep?*



Mother. *That's right darling, the doggy is asleep. Very tired! You want to sleep too?*

Child. *Baby sleep!*

Mother. *That's right, baby can sleep too! Come, mummy helps.*

Motherese contains many imperatives and questions, uses of language that require active involvement of the listener in the exchange. Other typical characteristics of motherese include high-pitched voice and profuse repetition.

The apparent idle play of children has a crucial role in language acquisition too. To give but one example, children who suddenly discover the thrills of playing a game like *peekaboo*, which demonstrates the permanence of an object or a face despite concealment, are well on their way to understanding features of language such as **arbitrariness** (referents are independent of their names, or different languages have different names for the same referent) and **displacement** (we can talk about things that are not present).

### Activity 3.2

Children's first words refer to people, objects and happenings in their immediate surroundings, e.g. spoon, bath, sleep, bottle, give, nappy, and, of course, mummy and daddy. Can you find reasons why this should be so?

Insight into the process of language acquisition, or **ontogenesis**, gained through intensive research since the mid 1950s, has renewed interest into the question of **phylogenesis**, or the origin of language itself. The question is whether ontogenesis can be said to replicate phylogenesis, and thereby help shed light into the age-old question

of how human beings came to develop language. Parallels that can be drawn between the patterns of early child speech and the most common patterns found in known languages appear promising. For example, early babble consists of repetitions of syllables of the form CV, or consonant followed by vowel, before children go on to tackle CVC, VC or other syllable shapes. Many languages have CV-shaped syllables only, and most languages that have other types of syllables have CV syllables too: a **CV-syllable** appears then to constitute a primeval component of words.

### 3.2.2 Language loss

The term **language loss** is usually associated with the waning or dissolution of language that concerns an individual speaker. Language loss can be caused by social factors like lack of prestige of a particular language, or language variety, due to value judgements associated with those languages, or to deliberate governmental policies. One example is the typical loss of the native language of second-generation immigrants, through pressure from peer or official environments where use of the native language is seen as refusal to conform to, or assimilate with, the mainstream or dominant culture.

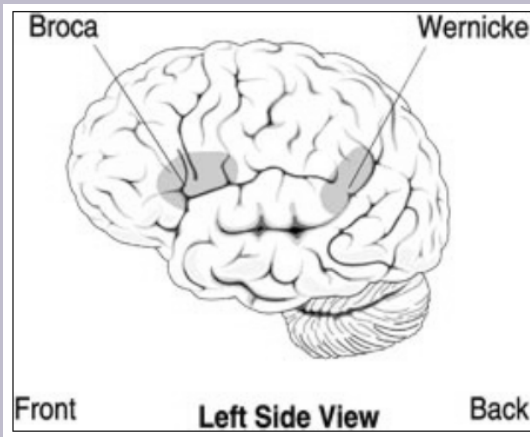
Language loss can also be caused by factors such as disease or trauma. The term for language disorders stemming from brain damage caused by physical injury or disease is **aphasia**. In the 1940s, Roman Jakobson (1941/1968) proposed that the patterns in aphasic loss of speech sounds mirror, in reverse order, those found in the typical development of language in children. That is, the first speech sounds to be acquired are the last ones to go. For example, **plosives** are among the first sounds to be acquired by children, and among the last to persevere in aphasia. Jakobson's interpretation of these observations as universal traits in language emergence and dissolution continues to raise controversy today.

Patterns in language pathology contribute insight to our understanding of human language in two chief ways. First, different modes of linguistic disruption test the robustness of the rules proposed to account for observed linguistic patterns, much like computer glitches test the robustness of a programme devised to perform a particular function. For example, if pathological conditions are found to result in the inability to use verbs, or inflected words, or [+ stop] sounds, then there is reason to believe that the word class verb, as well as the concepts of inflection and stop indeed constitute relevant theoretical constructs.

Second, disruption allows the setting up of hypotheses correlating particular types of **linguistic impairment** with specific locations of **brain** lesions. In the second half of the 19th century, two areas in the left hemisphere of the brain were found to play role in the production and in the comprehension of speech, respectively. The areas are named **Broca's area** and **Wernicke's area** (see Figure 3.1 below), after the researchers who first established that damage to these areas appear to result in particular types of speech impairment. Aphasics with injury to Wernicke's area, for example, produce fluent, grammatical speech whose lexical content is nonsensical. They also have difficulty understanding speech. In contrast, individuals with Broca's aphasia may have laboured speech, unusual word orders, and difficulties with function words such as to and if. Findings such as these suggest that grammatical and lexical processing of speech proceed along independent neural paths, and have spawned a flurry of current research into neural networks with the help of techniques such as functional neuroimaging.



Figure 3.1.



Approximate locations of Broca's and Wernicke's areas (With permission, from <http://www.nidcd.nih.gov/health/voice/aphasia.asp>)

### Activity 3.3

Here's one example of aphasic speech (from Goodglass and Geschwind, 1976). Well this is... mother is away here working her work out of here to get her better, but when she's looking in the other part. Would you say that this is more likely to be an example of Broca's or Wernicke's type of aphasia? Why?

### 3.2.3 Language death

Whereas language loss concerns individual speakers, the term **language death** is reserved for the extinction of a language affecting a community of speakers. Like language loss, language death can be caused by different socio- political factors. Regulating the use of language within national boundaries continues to be an effective means of controlling ideological dissent or access to power, on the well-founded premise that a language encapsulates the culture and values of its speakers. As the Portuguese poet Fernando Pessoa once said, “My motherland is my language”.

Policies of linguistic subjugation (or “unification”, or “planning”, depending on one’s point of view) are what banned the use of Catalan and Basque in General Franco’s Spain, and what lies behind debates that regularly flare up in multilingual countries like Canada and Belgium. **Minority languages**, or otherwise non-standard languages, are the usual targets of such policies. For example, banning the use of one minority language in schools effectively results in forcing its speakers to adopt the mainstream language, along with its culture and values. Monolingual speakers of the minority language are thereby barred from positions of power, for which official educational credentials are required. In practice, continued enforcement of such policies may result in the eradication of the targeted language from the country in question. If that language is not spoken elsewhere and, therefore, no new generation acquires it, the language effectively dies. Latin is often mentioned as the classic example of a **dead language** in that it is no longer transmitted across generations of speakers.

The global use of certain languages may also result in language death, this time because of suicide: speakers may voluntarily decide to stop using a language that they view as a hindrance to participation in a global community that uses another language. Linguistic globalisation therefore raises the parallel issue of **language endangerment**. The reason why so much attention is

currently paid to the preservation, or at least recording, of endangered languages is similar to that behind efforts to preserve and map the rainforest. Endangered languages are often spoken in remote parts of the world, where they remain untouched by global linguistic trends. Just as with the rainforest, there may be something there that tells us something we need to know, in that any language may provide us with invaluable insight about the nature of language itself.

Monolingual speakers of the three current **global languages**, English, Mandarin and Spanish, may count themselves lucky. By the happy accidents of birthplace or upbringing, these speakers have been raised to the enviable position of users of prestige languages. They need not worry about learning another language in order to be able to partake of the global cake. And, for the time being at least, the fate of Latin, a one-time global language too, need not worry them either.

#### Activity 3.4

See if you can find out which factors contributed to the death of Latin. Is there a cautionary tale here for today's global languages?

### 3.3 The multilinguals

As mentioned above, we use the term **multilingual** to refer to users and users of more than one language, regardless of the number of languages involved in multilingualism. We do this on the assumption that there may be a difference between the use of just one language

(monolingualism) and the use of more than one, but not between the use of two languages (bilingualism) and the use of more than two (multilingualism). In the literature, bilingualism is generally treated as being essentially different from monolingualism. To put it another way, the difference between monolingualism and bilingualism is seen as a difference in kind, whereas the difference between bilingualism and multilingualism is taken as a matter of degree.

One reason for this assumption may lie in the nebulous definition of **multilingualism** itself, and therefore of a multilingual. The term multilingual is used to label individual speakers as well as countries, individuals or groups of people that acquire several languages from birth, as well as those who learn a new language through schooling, or through settling in a different country. Clearly, each one of these is a different “multilingual”, although findings about each multilingual type tend to be generalised to multilinguals as a whole, by the use of the same word to label all users of more than one language. One added complication to the controversial definition of multilingualism is that there seems to be a reluctance to accept a multilingual as a **native speaker** of more than one language.

Definitions of a multilingual speaker range between extremes like ‘a multilingual knows several languages’ and ‘a multilingual is able to use several languages equally fluently in all circumstances’, both stumbling on the problem of how to quantify variables like “knowing” or “being fluent”, in order to draw comparisons. In addition, the latter definition begs the question: why would multilinguals need several languages, if they can do exactly the same thing with all of them?

We may want to start asking questions the other way around, in order to try to understand multilingualism. For example, why is a monolingual monolingual? Any monolingual will answer that they speak one language because they don’t need to speak more. The parallel with multilinguals and multilingualism then becomes clear – and perhaps not so odd, after all. People speak exactly the

number of languages that they need to speak, in different settings, to different people, and for different purposes.

### Activity 3.5

Work out a “map” of your own multilingual uses of language. Which language(s) do you use to whom, when, where and why? Discuss your map with a partner/partners, and compare your findings. Are all of you multilingual in the same way?

The fact that multilinguals have several languages at their disposal results in a sort of “buffet-effect” in their speech production, usually termed **mixes**. A mix concerns the occurrence of features that are ascribable to several languages in one utterance, and may involve any linguistic unit, from sounds through words to phrases. Just as a guest facing a rich gastronomic choice may want to sample the salad intended for the fish with a meat course, so multilinguals draw on the whole array of linguistic choices available to them in order to get their message through. Multilinguals do mix, but in exchanges with other multilinguals whom they know or suspect to share the same languages. In exchanges with monolinguals, multilinguals obviously recognise implicitly that mixes will result in disruption. Mixes do not, therefore, define multilinguals: they are simply the one feature of multilingual speech that arouses the curiosity of researchers, because it is not found in the speech of monolinguals.

Multilingual mixes are often discussed as evidence of poor command of language. Consequently, mixers are sometimes viewed as **semilingual**. One reason for assigning this special linguistic status to multilingualism lies in the fact that many linguists are monolinguals and/or subscribe to theoretical frameworks that were



devised to account for monolingual uses of language. Needless to say, trying to account for multilingualism from a monolingual perspective is rather like trying to understand siblinghood on the basis of one's experiences as an only child. Moreover, in terms of sheer number of speakers, multilinguals outnumber monolinguals, given that the majority of the world's population makes regular use of more than one language.

The view of mixing as a deficient use of language has deep historical roots that grow back at least to Ancient Greek thought, where language impurity was equated with mixing and change (anyone whose speech was unintelligible to monolingual educated Greeks was considered a "barbarian"). Here lies perhaps another explanation behind monolingual production being treated as the linguistic norm: the one language of monolinguals is treated as a language in its pure, unadulterated state, and therefore a true reflection of the human capacity for language. In view of our discussion, in section 1.1, concerning the ambiguity of the word "language" in the current language of science, English, such mix-ups (pun intended) are perhaps unsurprising. More importantly, they are reflected in virtually all the literature on mixing, where one language is taken as the core language of an utterance, upon which the other language(s) in the mixed utterance intrude(s). In this view, one language is seen to be disrupted by the other(s). If, on the other hand, we take mixed utterances as evidence of the use of language and not of the use of several languages, we may reach a different conclusion. Speech that, from a monolingual's perspective, is taken as mixed, may reflect instead the result of exploration of the accidental limits within which each particular language happens to vary, an exploration that is sanctioned by the open-ended nature of the language capacity itself.

An example may help clarify what we mean. The Malay word *malu* roughly means 'bashful'. In Singlish, a colloquial language variety in Singapore, utterances like *Very maluating* and *Very maluated* are attested to mean, roughly, 'very embarrassing' and 'very embarrassed'. Speakers of Malay may cringe at this defacing of a

word in “their” language: the original *malu* has not only been converted from an adjective to a verb, but has also been suffixed with foreign inflections. Speakers of English may in turn cringe at the intrusion of what clearly is a foreign verb stem, whose meaning they may not understand, into “their” language. For these speakers, the Singlish utterances above are mixed, neither Malay nor English, because they fail to follow the rules of Malay and of English. But language has no nationality, and therefore no owners, and its rules need not coincide with the rules of any individual language. If inflections of a particular kind are found useful in one language, why not overgeneralise them to another language where they happen not to exist? The plural of the “English” word *pizza* is *pizzas*, with an English inflectional suffix *-s*, not *pizze* as in the original Italian. Is the plural word *pizzas* a mix, then? From this perspective, it may well turn out that multilinguals do not mix at all. Rather, they are putting to communicative use the open-ended resources of language that are available to them.

### Activity 3.6

Find examples of mixes in your own speech, and in the speech of other multilinguals around you. Try to classify these mixes into categories like: Phonetic mixes: do you pronounce the words of one language with the accent and/or the intonation of another language? Lexical mixes: which lexical words do you find handier to mix into another language? Morphological mixes: do you add morphemes of one language to stems of another language? Syntactic mixes: do you use the words of one language with the syntax of another language? Discuss any patterns that you find with a partner/partners.

### 3.4 The others

The vagueness of the word “others” in this section is deliberate, and a form of self-inflicted linguistic **taboo**. All cultures have taboos, social bans restricting or prohibiting certain behaviours, which, if ignored, can result in social sanctions of various kinds.

The term itself is of Polynesian origin, and was first noted by Captain James Cook during his visit to Tonga in 1771. The Maori word *tapu* denotes the prohibition of an action or of the use of an object based on ritualistic distinctions between the sacred or consecrated, on the one hand, and the dangerous, unclean, and accursed, on the other. These social taboos often include linguistic taboos prohibiting the mention of certain events or entities considered either sacred (e.g. gods, religion, birth and death) or profane (certain bodily functions). Given the ban on words considered offensive in polite company, new words come to stand in for the tabooed ones, resulting in the occurrence of **euphemism** in all languages. Euphemistic words start off lacking the negative connotations associated with the tabooed words that they replace. But, because speakers know that euphemisms are stand-ins for tabooed expressions, over time the euphemisms themselves become negatively loaded and in need of replacement by new euphemisms.

The label “the others” is an instance of euphemism operating in the language of language. In what follows, we use it to refer to ‘non-native’ users of language. There are several reasons for avoiding the label non-native, one of them being that if there is no agreement on what a native is, then obviously there can be no agreement on what a non-native is either. The label **non- native speaker** generally refers to speakers who acquire a language either as a so-called second language, usually a language spoken in the country where the learner lives (for example, Malay in Singapore), or as a foreign language, where the language is not spoken in the country of the learner (for example, Japanese in Singapore). The

term “second language”, confusingly, applies even to cases where that language may be the learner’s third, or fourth, and so on. Both second language and foreign language learning situations typically involve traditional methods of language teaching usually in a school setting. A distinction is sometimes made between **language learning**, through schooling, as mentioned above, and **language acquisition**, through parent- child interaction.

A second reason for avoiding the label non-native is that, as we have observed in this chapter, where matters of language description encroach upon touchy human matters of culture and national policies, scientific labels may undergo the same fate as euphemisms, becoming loaded words instead. The fact is that both the word native and its presumed opposite non-native have acquired connotations that complicate their definition in any scientifically useful way. A molecule or a prefix won’t feel affronted by being called molecule and prefix, whereas human beings can and do take offence at being called non- native. In much current research, both words are simply replaced by euphemistic acronyms, NS for ‘native speakers’ and NNS for ‘non-native speakers’, with no attempt at defining these. Alternative labels include first- language learner vs. second-language learner, or L1 user vs. L2 user, none of which have been usefully defined either.

For example, should people who speak, from birth, a language that was once imported to their country be labelled native speakers of that language? This is the situation faced by many speakers of English in India and most speakers of Portuguese in Brazil. But the language situation in these two countries is quite different. In terms of official language policy, Brazil is a “monolingual” country, because it has one official language, whereas India is not a monolingual country, because it has more than one official language.

Another example of a touchy issue concerns the current debate about whether Spanish should be recognised as an official language in the United States, given the increasing weight of the language in the country. Granting official status to a language means of course that its speakers, including monolinguals, are to be treated on equal

footing with speakers of other official languages, for all purposes and in all circumstances. This is where language, ideology and power become enmeshed. There is a huge difference between labelling Spanish speakers as non-native speakers of the official language of their country and labelling them as native speakers of one of the official languages of their country. The former labels them as outsiders (the others of our section heading) whereas the latter empowers them as insiders.

### Activity 3.7

Re-read the last two sentences above, repeated here for convenience: “There is a huge difference between labelling Spanish speakers as non-native speakers of the official language of their country and labelling them as native speakers of one of the official languages of their country. The former labels them as outsiders (the others of our section heading) whereas the latter empowers them as insiders.” Do you agree that the difference we are talking about here is “huge”, whether for Spanish speakers in the United States or for other speakers in a similar situation? Why? In what way are insiders “empowered” through the use of one language?

In much of the literature, non-native uses of language are described as instances of **multilingualism**, in the way suggested in section 3.1, particularly where use of the dominant language of a country by immigrant populations is concerned. Whatever the definition of multilingualism, it clearly concerns **language contact**, including in its pidgin and creole forms. As was also remarked there, contact uses of language provide valuable insight into language in the

making. For example, multilingual children and pidgin speakers often use the grammatical constructions of one of their languages with the lexical words of the other(s), pointing to separate neural processing of the two levels of linguistic structuring. The same separation was noted in instances of language dissolution, in section 3.2.2.

In addition, instances of multilingualism constitute strong factors of **language change**, in that the users, having access to more than one language, are at greater freedom to explore the creativity of language itself. Multilingual exploration may proceed through **overgeneralisation** of perceived rules, mirroring the common process in language acquisition mentioned in section 3.2.1. Children's use of language constitutes another important factor of language change. Drawing on the morphological pattern of words like *cooker* and *blender*, monolingual children as well as multilinguals of all ages may produce forms like *clipper* for 'scissors' or *pumper* for 'pump', usefully compositional forms which may be "wrong" from the perspective of common uses of English but are certainly "right" from the perspective of possible uses of the language. Any word or construction that is current in a language must obviously have been introduced sometime in the history of that language in precisely this novel way, and thereafter gained use through acceptance of its usefulness by other language users.

### Activity 3.8

The following sentence was said by a non-native speaker of English. The underlined word in the sentence is not a standard word of English. We have been looking for those weapons, but the search has been resultless so far. Can you explain how the speaker created this novel word? Can you

explain why the meaning of the word is clear to speakers of English?

Unexpected uses of language such as these raise the matter of **intelligibility**. In linguistic exchanges involving one language, the common assumption among speakers seems to be that if we speak the same language, then we do it in the same way. Recall, from section 1.1, that sender and receiver must share the ability to decode each other's message, so that communication can take place. "One language" is thus taken to mean 'the same code'. In actual fact, the situation is quite different. Take the case of communication in English. Its status as a global **lingua franca** means that most of the exchanges in English around the world take place among native speakers of other languages. A Portuguese businessman attempting a deal with a Japanese counterpart will in all likelihood speak the English that he learned in school, in a Portuguese classroom, from native speakers of Portuguese, and with the help of Portuguese glosses and paraphrases to clarify obscure uses of English. That is, he will have learnt English not as a language in itself but as some "variant" of Portuguese. The same is true of the Japanese speaker of English, both speakers being unaware that they are in fact speaking different Englishes, and that disruption may therefore arise in their uses of their "shared" language.

Here's one example that one of the authors of this book witnessed at an international conference. An Asian participant gently reminded the Scandinavian presenter that she had exceeded her allotted twenty minutes, and asked if he could please ask some questions on her very interesting paper. The presenter checked her watch, apparently baffled by her miscalculation. Turning to the Asian gentleman, she cried "It's not true!", a literal translation of a Scandinavian apology rendered into English. The reaction of the Asian gentleman, and of most of the remaining audience, was to leave the room. The Scandinavian speaker had unwittingly insulted

her audience by implying that they were liars. The irony is that the conference was on the topic of teaching English as a second language.

Proficiency differences between first and second/foreign languages, ranging from accent to pragmatic uses, have fed the much-debated issue of the so-called “**critical period hypothesis**” – the belief that the human capacity for learning language is limited to a critical age-range, variously set between early childhood and the late teens, beyond which the acquisition of a new language is either impossible or severely impaired. As evidence, researchers point to the failure of non-native speakers to reach native-like linguistic proficiency, a goal that is in itself questionable: which native variety should learners strive to emulate, and for what purposes? But the main issue is that supporters of a “critical period” fail to take into account the ways in which the new language is learnt. In traditional school settings, for example, learners are force-fed vocabulary lists and rules of “grammar” instead of being given the chance to use the language naturally in a variety of settings. Expecting these learners to achieve full linguistic proficiency is like giving aspiring bike-riders a description of the component parts and mechanics of a bicycle, and expecting them to be able to ride it. Fluent command of language arises from natural interaction among speakers, the one form of learning to which second and foreign language users generally have little or no access. In addition, keeping in mind that motivation, not age, is the prime mover of human achievement, including language acquisition, the assumption of a “critical learning period” is in fact a non-issue.

Matters of intelligibility arise not only across languages, but also within languages. Users of different language varieties might be perceived as non-native speakers by those who speak a different variety of the same language. In Chapter 2, we highlighted the twin phenomena of linguistic **convergence** and **divergence**. In linguistic convergence, speakers adapt their speech patterns at the level of word, grammar and/or intonational choices to speak more like their conversational partners, thereby narrowing the sociolinguistic



differences between themselves and their partners. We can see convergence at work in the way adults accommodate their speech when addressing young children, in order to match the child's linguistic proficiency. Similarly, when speaking to someone perceived to be of lower status than ourselves, we converge towards their speech patterns in order to reduce social distance through speech. Conversely, speakers may choose linguistic divergence to highlight the social differences between themselves and their conversational partners.

### Activity 3.9

Find patterns of convergence and divergence in your own uses of language. Try also to work out the reasons for your choices of language use when you converge with, or diverge from, your interlocutors' use of language.

In terms of speakers of different varieties of the same language, the question that arises is: do non-natives want to speak like natives and, if so, which natives? Research into these matters reveals apparently paradoxical findings, to the effect that a British accent, say, is evaluated by non-British listeners as correlating with higher levels of intelligence, education and politeness compared to other accents of English, while the same listeners state that they would at all costs avoid the use of such an accent because they don't want to sound "posh" or "pretentious" (recall the Lette quotation in the Food for thought section of Chapter 2). Faced with linguistic dilemmas such as these, speakers often settle for what can be usefully viewed as another form of multilingualism, where international-like and local-like varieties of the same language are used in distinct situations. Examples of these choices are the uses to which

Singaporean speakers put varieties like Singapore Educated English and Singapore Colloquial English (Singlish), or the uses of Standard German and Swiss German (*Schweizerdeutsch/Schwyzerdütsch*) in German-speaking parts of Switzerland.

In our discussion of dialects, we highlighted that a dialect is a regional language variety which characterises a speech community, whose members *choose to see themselves* as speakers of the same language. In other words, whether or not some variety is a language or a dialect is as much a socio-political question as it is a linguistic one. The same comment can be made about intelligibility. Whether two speakers find themselves mutually intelligible has to do with the relationship among them. Intelligibility concerns whom we are communicating with, whether we really want them to understand us, and whether they really want to understand us. Local uses of a language, from accent through syntax to pragmatics, can effectively screen off uninitiated speakers, and therefore be used as a weapon in demarcating one's territory, or linguistic identity. As one Hong Kong native once put it, in response to an Englishman's baffled query about whether it was really English that people were speaking to him, "Everybody speaks English in Hong Kong, but nobody understands what you say."

### 3.5 Several speakers, one language

Our investigation of human language and its users throughout this book may at first sight suggest that human beings all over the world are talking at cross-purposes from within the well-protected codes of their individual languages. A closer look, however, reveals the opposite trend. The recurrent and fruitful application of constructs like lexical word, constituent, phoneme, distribution, intonation, register, to forms and uses of language across different languages and speakers compels us to realise that we all speak the same language. This is no different from saying that I am similar to a

gecko or a lion in terms of my anatomical structure in that all three of us can be usefully described by a label like vertebrate, which distinguishes us from invertebrates like jellyfish and snails.

The lexicon of linguistics, dreaded by generations of budding linguists, is no different either from the lexicon of any other language, technical or otherwise. In the same way that it is easier to use a word like cat for the complex being designated by the term cat, it is also more economical to use shorthand like suffix for ‘a bound morpheme that attaches to the right of a stem’ and homophones for ‘words that sound the same but are spelt differently’. Learning technical terminology is like learning a foreign language, at times more puzzling because the words of that language may be the same as those of familiar languages, only with new meanings. The language of language, by the very nature of its object and its users, encapsulates a culture, too, with associated taboos, ambiguities and deliberate vagueness, and with a choice of labels that, expectedly, reflects human interaction and human forms of social organisation. Natural classes, contexts and alternations mirror peer-groups, favourite hangouts and variant ways of behaving in different settings in everyday life. Similarly, linguistic heads, sisters and adjuncts behave in similar ways to the human beings in the relationships described by these labels.

**Linguistics** provides us with the tools that crack the code of our common language. Giving you a first glimpse into the unifying nature of the language of language has been the purpose of this book.

## Food for thought

Emir (age 4): “I can speak Hebrew and English.”

Danielle (age 5):

“What’s English?”

Quoted in Jill G. de Villiers and Peter A. de Villiers

(1978).

*Language acquisition*. Cambridge MA: Harvard

University Press.

“And what should they know of English who only English know?”

Adapted from

Rudyard Kipling (1891/

1949). *The English Flag*.

In *Rudyard Kipling’s*

*Verse*. London: Hodder

and Stoughton.

Browse through this book and through the readings suggested in each chapter, this time from the perspective of linguistic forms and uses that you’re familiar with from languages other than English, the language from which most of our examples were drawn. In doing so, make it clear to yourself why the shared patterns found in different languages indeed support a view of spoken language as a cohesive whole.

## Further reading

Aitchison, Jean (2001). Chapter 10. The reason why: Sociolinguistic causes of change. In *Language change: Progress or decay?* (3rd ed.). Cambridge: Cambridge University Press, pp. 133-152.

Gleason, Jean Berko (2005). Chapter 1. The development of

language: An overview and a preview. In Jean Berko Gleason (ed.), *The development of language* (6th ed.). Boston: Allyn and Bacon, pp. 1-38.

Kachru, Braj B. (ed.) (1992). *The other tongue: English across cultures* (2nd ed.). Urbana: University of Illinois Press.

Obler, Loraine K. (2005). Chapter 11. Developments in the adult years. In Jean Berko Gleason (ed.), *The development of language* (6th ed.). Boston: Allyn and Bacon, pp. 444-475.

## References

Goodglass, Harold and Geschwind, Norman (1976). Language disorders. Aphasia. In Edward Carterette and Morton P. Friedman (eds.), *Handbook of perception: Language and speech* (vol. 7). New York: Academic Press.

Jakobson, Roman (1941/1968). *Child language, aphasia, and phonological universals*. The Hague: Mouton.

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.

# Chapter 4: Phonology

## Speech sounds

### Chapter Preview

How are speech sounds different from other sounds?

How do we articulate speech sounds?

What kind of sounds are used in languages?

### 4.1 Introduction

Any sound, whether used in speech or otherwise, can be characterised in terms of its quality. **Sound quality** refers to a specific set of acoustic properties that distinguishes one sound from another.

Being about language, this book is concerned with **speech sounds**, sounds produced by human beings for the purposes of linguistic

communication. This definition encompasses the two branches into which the study of speech sounds divides:

- **Phonetics** deals with the articulatory capabilities of the vocal tract, and therefore with an ability shared by all human beings. These capabilities are the object of *articulatory phonetics*, the branch of phonetics that we deal with in this chapter. Articulatory phonetics studies how speech sounds are produced (articulated) in the human vocal tract. *Auditory phonetics* and *acoustic phonetics* constitute further sub-branches of phonetics, not dealt with in this book. Auditory phonetics studies the mechanisms involved in speech audition, i.e. how listeners perceive speech sounds, while acoustic phonetics studies the physical characteristics of speech sounds.
- **Phonology** deals with sounds that serve linguistic purposes. These are sounds whose patterning conveys meaning. These sounds are recognised by speakers and listeners as linguistic signals in the language(s) that are available to them, and are therefore shared by speakers who share a language. We will look more closely at phonology in the next chapter.

In talking about sounds, we need to keep in mind that there is often no one- to-one correspondence between speech sounds and their spellings. So, we need to (re)train ourselves to listen to speech instead of reading printed forms of it. Here are a few examples of sound-spelling discrepancies in English. The words *threw* and *through* are pronounced in exactly the same way, as are *knows* and *nose*: when spoken out of context, we can't tell which is "witch". Conversely, the English letter sequence *-ough* has several different pronunciations, as in words like *cough*, *dough* or *through*, and the sound [k] can be spelt in at least eight different ways, as shown by the bolded letters of the following words:

(5.1) *tack***k** , *cat*, *mech***an***ic*, *squid*, *bea***k** , *ac***q***uire*, *ac***c***ordion*,  
*grotes***q***ue*

A few more examples from English appear in the Food for thought section at the end of this chapter.

#### Activity 4.1

Can you find examples of discrepancies between speech sounds and their spellings in other languages that you are familiar with?

## 4.2 The production of speech sounds

In order to communicate through speech, human beings move the lower part of their faces, in a frenzy of rapidly shifting and very precise configurations. If you think of yourself as a lazy or slurred speaker, and would dismiss the word *precise* as inapplicable to you, just take some time to observe the tremendous effort that small children put into the task of training the muscles that command the production of speech: you also had to go through this painstaking workout of speech musculature, in order to be able to “slur” fluently.

The lower jaw, being mobile, stands for the visible part of speech configurations. The role of the lips in speech is clearly visible too, as is, at times, the configuration of the tip of the tongue and the teeth in the articulation of certain speech sounds. However, most speech configurations take place inside the oral cavity, anywhere along the vocal tract. They are therefore difficult to see, but fairly easy to feel. The bulk of this chapter is dedicated to showing you what it takes, and how it feels, to pronounce a range of common speech sounds.



## 4.2.1 The vocal tract

The **vocal tract** forms an inverted L-shaped tube that stretches from the larynx upwards to the lips. The angle of the L is at the back of your throat. The **larynx**, or the *voice box*, is located mid-way down the front of your neck. If you're male, you may find that the so-called "Adam's apple" of your larynx can be quite protruding. The vocal tract includes the nasal cavities, which are located inside and behind your nose, as optional resonator.

We normally speak while breathing out (you can try uttering a few words, or maintaining a conversation while breathing in, to get a feeling of how unnatural this is). This explains why speaking for long periods of time can be quite tiring. The repeated interruptions of the airflow that make speech possible disrupt the normal rhythm of breathing. Speech sounds are produced by means of modifications imposed on exhaled air, as it flows through the vocal tract. These modifications are brought about by movement of the **articulators**, each of the component organs and locations in the vocal tract that play a role in the production of speech sounds. Generally, the **active articulators** along the lower jaw move towards the passive articulators in the upper jaw. Both lips, lower and upper, can be active articulators.

Instead of providing you with the usual labelled diagram of the vocal tract found in most introductory material on phonetics, we offer in this chapter "hands-on" acquaintance with the basic anatomy and physiology of your oral cavity. We will show you around your own vocal tract, making it clear to you what goes on inside it when you speak. In the process, the reason why so many apparently obscure labels are needed for the description of speech sounds will, we hope, become clear to you. As a preliminary observation, keep in mind that just as we need different labels for words that behave differently, so we also need different labels for each of the vocal movements that produce a different speech sound.

In what follows, we provide you with a guided tour of your vocal

tract, together with core speech configurations used in many languages, including English. In order to fully understand this chapter, you will need a mirror, a pencil or pen, and some privacy. Warn anyone near you not to panic at the noises you'll be producing: phonetics must be practised, it cannot be learned otherwise. Better still, work through this chapter with a friend or two, willing to observe and be observed, and practise together! Just follow the Try the following trail.

## 4.2.2 Core speech configurations

Modifications along the length of the vocal tract affect its shape, and therefore its volume and resonance. Each modification thus produces a different sound. In this book, we deal with only a few of these modifications, most of which apply to most languages, including English. According to these modifications, speech sounds may be classified as *voiced* or *voiceless*, *oral* or *nasal*, and as *vowels* or *consonants*. We deal with each of these alternative articulations in turn.

### *Voiced vs. voiceless sounds*

Try the following

Gently touch the section of your larynx that slightly protrudes at the front of your neck. Now say something like *Good morning!*, loud and clear. Can you feel the vibration under your fingers? This is the effect of your **vocal cords** at work, two folds of mucous tissue inside your larynx, that vibrate in order to produce **voice**. Now, still keeping your fingers in place, whisper *Good morning!* You'll notice that the vibration is gone, because

whisper has no voice. Your vocal cords do not vibrate when you whisper. When we produce speech sounds, the vocal cords can be in one of these two states:

- *Closed*, i.e. loosely brought together so that they vibrate when air flows through them. It is this vibration that creates what we call voice, resulting in **voiced** sounds.
- *Open*, i.e. pulled apart during the production of certain sounds, as they are during normal breathing or in whisper, resulting in **voiceless** sounds.

Try the following

With your fingers again in place on your larynx, say a long *zzzz* as in the beginning of the word *zap* and a long *vvvv* as in the beginning of the word *vat*, loud and clear. Next, say a long *ssss* (*sap*) and *ffff* (*fat*) sound. Did you notice that the first two sounds are voiced and the last two are voiceless? For a more dramatic contrast, say the same sounds with your hands over your ears, blocking them off. The two voiced sounds *zzzz* and *vvvv* produce a buzzing vibration, caused by voicing, inside your head, whereas the two voiceless sounds *ssss* and *ffff* produce a turbulent hiss of air.

### *Oral vs. nasal sounds*

Try the following

With the tip of your tongue, touch the roof of your mouth. You'll feel a bony surface, called the **hard palate**. Now drag the tip of your tongue back along the hard palate, as far back as it will go. You will notice that at the very back there is no bone:

this is the **soft palate**, or **velum**.

Now close your lips tightly, and hum. Notice that you don't need to use your jaw at all to make your voice heard. Sounds that are produced as you hum involve using your nasal cavity: you produce them as air flows out through your nose. When we produce speech sounds, the soft palate can be in one of these two states:

- *Raised* against the top part of your pharynx, which is the back wall of your throat. When raised, the soft palate blocks the airflow to the nasal cavities, resulting in **oral** sounds.
- *Lowered*, causing the air to flow through the nasal cavities, as when you hum, resulting in **nasal** sounds.

Try the following

Say a long *ah* sound, as if calling for help by using the word *Guard!* Now, say the sound given in English spelling by *-ng*, as in the word *bang*. Repeat the sequence *ah-ng-ah-ng-ah-ng-ah* several times with no pauses in between, paying attention to what's going on at the back of your mouth. You will feel the soft palate moving up and down, up for *ah*, which is an oral sound, and down for the nasal sound *-ng*. While doing this, pinch your nostrils together. You'll have no problem producing the oral sound *ah*. But you're no longer able to produce the nasal sound *-ng*, right? This is because when you pinch your nostrils together, air cannot flow out through your nose. To double-check that it is your soft palate moving, try saying *ah-gah-gah-gah*, where both sounds are oral. Pinch your nostrils together again, to confirm the difference between oral sounds like *gah* and nasal ones like *ng*. Assuming that you've had a bad cold at least once in your life, you'll remember the strange

resonance of your speech at the peak of it, when you have a blocked nose. Your listeners notice it, too. What happens is that air cannot flow out through your nasal cavities to produce the nasal sounds of speech. You may now realise that the popular expression “speaking through your nose” couldn’t be further from the facts: with a bad cold, you speak exclusively through your mouth.

#### Activity 4.2

Pinch your nostrils together again to replicate a really bad cold, and complain about your symptoms:

*My, **my**! I feel all numb and funny!*

Can you see that what you’re really saying is something like Bye, bye! I feel all dub add “fuddy”? What can you conclude about the articulation of the sounds represented by the bolded letters in the words *my*, *numb*, *and*, *funny*, in the original utterance?

#### *Vowels vs. consonants*

Try the following

Say the long *ah* sound again. You’ll notice that not much is

happening inside your mouth when you say it. You simply open your mouth and produce voice.

Now place the tip of your tongue against the slight protrusion behind your upper front teeth, where your gums start. This is called the **alveolar ridge**. Say the sound *t*. Did you notice that in order to produce this sound, the tip of your tongue is forced away from the alveolar ridge? This type of sound requires serious involvement of the articulators.

Generally, we can say that:

- **Consonants** like *t* are produced by means of an obstruction to the airflow in the oral cavity. The lips, teeth and tongue form major consonantal obstructions.
- **Vowels** like *ah* are produced with no obstruction in the mouth: the shape of the vocal tract is modulated by different configurations of the tongue and lips.

Note that vowels and consonants refer to sounds, not letters/ spellings. If you think there are five vowels in English, *a, e, i, o, u*, you will have reason to change your mind after you read the next chapter.

#### Activity 4.3

Say a long *oh*, and then a long *sh*! (as if shushing someone).

Which one is the vowel and which is the consonant?

### 4.2.3 Consonants

We can produce different types of consonants, depending on two major factors. One is the degree of obstruction to the airflow, which gives us the **manner of articulation** of the consonant. The other is the location of the obstruction along the vocal tract, which tells us about the **place of articulation** of the consonant. These two factors operate independently of each other: in the same way that you can place obstacles of the same or different kinds just about anywhere along a cross-country obstacle course, you can have the same or different types of obstructions to the airflow in different places along your vocal tract.

#### *Manner of articulation*

- **Plosives** are pronounced with the velum raised, and involve contact of articulators. This means that the articulators first touch each other, and then separate.
- Non-plosive consonants, in contrast, involve close approximation of the articulators. This means that the articulators come very close to each other, without ever touching. One example is the class of **fricative** sounds, where the air is hissed (i.e. expelled with *friction*, hence their name) through a very narrow gap between the articulators.

Try the following

Check that the t-like sound that you produced before is a plosive: the tongue tip touches the alveolar ridge, thereby closing off oral airflow. The sudden release of this closure produces an *explosive* effect.

Try now the long sounds *ssss* and *vvvv* again. The friction that you hear means that there can be no contact between articulators: there must be a narrow channel somewhere that allows the air to hiss through it. Note that you can prolong fricative sounds for as long as you have air in your lungs. This is not possible with plosives, whose articulation is instantaneous. To check, try prolonging a *t* sound – chances are that you will go red in the face for lack of air, because of the plosive closure that is needed to produce this sound.

### *Place of articulation*

- **(Bi)labials** involve one or both lips in their articulation.
- **Alveolars** involve the alveolar ridge as articulator.
- **Velars** involve the soft palate/velum as articulator.

Try the following

Now you know that *t* is not only plosive, it is also alveolar. You may also have guessed by now that the sounds spelt *p* and *b* are bilabials, involving both lips. Use your mirror (or your friends) to check these clearly visible articulations.

Say the sequence *ah-ng-ah-ng-ah-ng-ah* again, this time to check that the *-ng* sound involves contact between the back part of your tongue and the velum. So do the initial sounds in the words *gap* and *cat*. You can check this by saying *ah-gah-gah-gah* and *ah-kah-kah-kah*. All three sounds are therefore velar.

Now say *vvvv* and *ffff*, and concentrate on which articulators are



involved in saying them. Use your mirror-friend to check. The upper lip plays no role in their articulation, but the lower lip does. These sounds are pronounced with the upper teeth very close to the lower lip. These sounds are therefore labial (and *dental* too: they involve the teeth, that your *dentist* checks out for you).

The place of articulation of sounds like *ssss* and *zzzz* is more difficult to check, because the whole of their articulation goes on inside the mouth. In terms of manner of articulation, both sounds are fricative. There is therefore no contact of articulators that may help us feel the place of articulation, in the absence of visual clues. But try this: say *ssss* and *zzzz*, then stop saying them but keep the articulators in exactly the same position, and breathe in quickly. You will feel a rush of cold air between the tip of your tongue and the alveolar ridge. This is the space inside your mouth that allows the articulation of the sounds *ssss* and *zzzz*, which shows that the constriction to the flow of air is located there. Both sounds are therefore alveolar.

#### 4.2.4 Vowels

The articulation of different types of vowels depends on the movements of the tongue and lips. The body of the tongue can move *vertically* (up or down) and *horizontally* (front or back). The lips may be involved in the articulation of vowels, too.

The movements of the **tongue body** (the bulk of the tongue inside your mouth) are independent of one another. You can, for example, move your tongue up and down whilst bunching it up at the back of your mouth or, conversely, you can push it to the front or the back of your mouth while holding it high towards the hard palate. Try it.

Tongue-body movements are also independent of any movements of the **tongue tip**: the tongue tip is used in the production of consonants, not vowels. It may in fact come as surprise to you, especially if you are a body- builder, that the tongue is the best

trained and most versatile muscle in your body. Its core role in speech is clear from the use of its name, “tongues”, to mean ‘languages’, in many languages around the world. We don’t notice the spectacular gymnastics show that goes on inside our mouths every time we speak because we have been practising this sophisticated skill since infancy. To get an idea of the range and power of tongue actions, remember it is used in chewing and swallowing, too. If you don’t believe us, try chewing on something while keeping your tongue motionless!

### *Vowel height*

- **High** (or **close**) vowels are produced with the body of the tongue raised.
- **Low** (or **open**) vowels are produced with the body of the tongue lowered.

Try the following

Open your mouth wide and say a long aaaah again. While still saying ah, slowly raise your jaw. Use your mirror to check that you are indeed raising your jaw. You’ll find that, whatever you’re saying now, it’s not ah anymore. The vowel ah, in a word like cart, must be pronounced with a lowered jaw, and therefore a lowered tongue: it is an open vowel.

Now say a long iiiih vowel, as in the word see, and drop your jaw while saying it. You won’t be able to say ih anymore. The vowel ih needs a raised tongue to be produced: it is a close vowel. The difference between open and close vowels explains why your doctor asks you to say aaaah, and not iiiih, in order to be able to examine your throat.

## *Vowel backness*

- **Front** vowels are produced with the body of the tongue pushed forward in the oral cavity.
- **Back** vowels are produced with the body of the tongue pulled back/retracted.

Try the following

Say the vowel in the word *cat*, and then the vowel in the word *cart*, several times in a row. You'll feel your tongue moving back and forth while doing this, back for the vowel in *cart*, and forwards for the vowel in *cat*. Use a mirror, to get visual feedback on this: it is easy to see because both these vowels are open.

Now try the same thing with the vowels in the words *bean* and *boon*. This is more difficult to feel because the *boon* vowel involves a distracting factor, which is the pouting, or rounding, of the lips. The difference in tongue movement is also impossible to see, because both vowels are close. So try this, instead. Take a pen, or a chopstick, and place it lengthways across your mouth, so that both ends of the pen/chopstick stick out from the sides of your mouth. Then bite it firmly as far back in your mouth as possible. Biting the pen/chopstick prevents the lips from moving in the articulation of the vowel in *boon*. Now say the vowels in *bean* and *boon* again. You'll feel your tongue touching the pen in the articulation of the vowel in *bean*, but not of the vowel in *boon*. We thus conclude that the vowels in *cat* and *bean* are front vowels, whereas those in *cart* and *boon* are back vowels.

## *Lip rounding*

- **Rounded** vowels are produced with rounded (pouting) lips.

- **Unrounded** vowels are produced with unrounded (spread) lips.

Try the following

Easiest comes last. Of the four vowels that we've discussed, the vowel in *boon* is the only one that involves the lips: the lips must be rounded in its articulation. This is also why your photographer asks you to say *cheeeese* and not *choooooose* when prompting a smile from you: spread-lip vowels are “smiley” sounds.

## 4.3 The transcription of speech sounds

A **phonetic transcription** is a way of representing sounds in print. The **IPA** (International Phonetic Alphabet) is a standard among **phonetic alphabets**, with symbols that correspond bi-uniquely to the sounds of the world's languages. This means that, unlike spelling symbols, the same phonetic symbol *always* represents the same sound, and vice versa. In other words, unlike the relationship between speech sounds and spelling, there is a one-to-one correspondence between phonetic symbols and the sounds they represent. This is why trained phoneticians can read a phonetic transcription of any language and sound fluent in it, even if they have no idea how to speak (or spell!) the language.

Figure 4.1 below gives the set of **phonetic symbols** used in this book. In the token words included in the table, the letters that correspond to the sounds on the left are in bold italic font. When consulting phonetics literature, or dictionaries that include IPA transcriptions, you may notice that certain vowels are transcribed using an additional symbol [ː] following the vowel symbol. This indicates vowel length, and a corresponding distinction between long and short vowels present in some languages and some varieties of English, e.g. *bean* vs. *bin*, respectively. Many dictionaries also

mark the stressed syllable of words. Intonation in turn has its own set of transcription symbols. For the purposes of this book, you do not need to be familiar with these conventions.

Phonetic symbol	Token word	Phonetic symbol	Token word
<b>p</b>	<b>pat</b>	<b>f</b>	<b>feel</b>
<b>b</b>	<b>bat</b>	<b>v</b>	<b>veal</b>
<b>t</b>	<b>team</b>	<b>s</b>	<b>seal</b>
<b>d</b>	<b>deem</b>	<b>z</b>	<b>zeal</b>
<b>k</b>	<b>card</b>	<b>i</b>	<b>bean</b>
<b>g</b>	<b>guard</b>	<b>æ</b>	<b>ban</b>
<b>m</b>	<b>mean</b>	<b>a</b>	<b>barn</b>
<b>n</b>	<b>seen</b>	<b>u</b>	<b>boon</b>
<b>ŋ</b>	<b>pang</b>		

Figure 4.1. Phonetic symbols used in this book

Phonetic transcriptions are usually given in square brackets. For example, the transcription of the word cat is represented [kæt].

Activity 4.4

Practise transcribing these words, using the phonetic symbols introduced above.

Remember to focus on how the words are pronounced, not on how they are spelt.

fa ng	f goo	mo ose	pi ece
sc arf	sn ack	spe ak	ta x

## 4.4 The analysis of speech sounds

According to the phonetics literature, there are two major ways of describing and labelling speech sounds:

- The **IPA** (International Phonetic Association) approach. Note that the abbreviation IPA stands for two different things – an alphabet and an association.
- The **DF** (Distinctive Feature) approach.

Both approaches draw on articulatory features of speech such as the ones that we have discussed in this chapter, and both therefore propose a universal framework for the description of human speech sounds. However, each approach takes a different perspective to look at speech sounds. This naturally results in alternative classifications of speech sounds, that do not always overlap across the two frameworks. As we've noted before, alternative analyses of this kind are as common in linguistics as in other sciences. They are part and parcel of our quest to understand, and they should be viewed as the probing tools that they are. It's up to us, as users

of these tools, to choose the one that we deem most useful to approach the object of our inquiry, or to come up with a more illuminating alternative.

There are two major differences between the IPA and DF frameworks.

- The IPA approach views vowels and consonants as fundamentally distinct types of sounds. The DF approach views all speech sounds as fundamentally similar.
- The IPA assumes a set of articulatory configurations whose combined presence in the articulation of a sound defines that sound. The DF approach assumes a set of articulatory configurations, each of which can be present [+] or absent [-] in the articulation of a sound, on a binary basis.

We now comment on each of these differences in turn. Because IPA analyses of speech sounds view vowels and consonants as radically different in terms of their articulation, they make a principled distinction between the classification of vowels and consonants. DF analyses don't.

The IPA uses two distinct sets of articulatory configurations, one to characterise vowels and another to characterise consonants, each with their own set of terms. Consonants are described in terms of three articulatory configurations, namely, vocal cord vibration, place of articulation, and manner of articulation. Any consonant can be uniquely described using these three criteria. For example, [d] is a voiced alveolar plosive; [s] is a voiceless alveolar fricative. Vowels are described in terms of three other articulatory configurations, namely, lip rounding, tongue height, and tongue backness. So the vowel in *bean*, for example, can be uniquely described as unrounded close front, while that in *boon* is rounded close back.

In contrast, the DF framework uses the same set of criteria for all speech sounds. Each configuration of the vocal tract corresponds to an articulatory feature that allows us to distinguish one speech sound from another. These features are therefore *distinctive*. For

example, raising the body of the tongue characterises the articulation of sounds like [i] and [k], as opposed to [ɑ] and [z]. A feature that allows us to distinguish between the first and the second set of sounds is therefore relevant for our characterisation of speech, and is called [high] (distinctive feature labels are usually represented in square brackets). Or, vocal cord vibration, labelled [voice], allows us to distinguish [i, ɑ, z] from [k]. Or, the smoothness of the airflow through the vocal tract distinguishes [i, ɑ] from [z, k], and is represented in a feature called [sonorant].

The second difference between IPA and DF analyses concerns the way each of them views the articulatory make-up of a speech sound. For the IPA, a sound consists of a pool of articulations that, together, produce that sound. For example, [d] is voiced, *and* alveolar, *and* plosive, and [i] is unrounded, *and* close, *and* front. To use an analogy, a girl would likewise be characterised by the pool of properties human, *and* female, *and* child. The *presence* of these features makes a sound (or a girl) unique. In DF approaches, a *distinct* sound depends on whether specific *features* of articulation are activated or not in its production. For example, [i] is [+sonorant] and [z] is [-sonorant], [d] is [+voice] and [t] is [-voice]. A girl could be described as [+human -male -adult]. The *interplay* of features, activated as well as inactivated, is unique to the sound.

Given the introductory nature of this book, we will not go into theoretical detail relating to the two frameworks. However, you will come across published material on phonetic analysis that often makes free use of terminology from either approach, and often in the same research piece. The set of tables that we propose below will, we hope, facilitate your learning about phonetics.

We start with the IPA, because this was the first analytical framework available to phoneticians. The IPA was founded in 1886, and the first DF proposals appeared in the early 1950s. Throughout the discussion in the remainder of this chapter, you should keep in mind that the use of the International Phonetic Alphabet in transcriptions does not imply subscription to the classificatory approach of the International Phonetic Association. DF analyses



may use IPA phonetic symbols. A set of printable symbols and a theoretical stance about the classification of the sounds that these symbols represent are two different things.

*IPA representations of speech sounds*

Below are two samples of IPA charts, one for consonants and one for vowels.

		Place of articulation		
		(bi)labial	alveolar	velar
Manner of articulation	plosive	<b>p b</b>	<b>t d</b>	<b>k ɡ</b>
	nasal	<b>m</b>	<b>n</b>	<b>ŋ</b>
	fricative	<b>f v</b>	<b>s z</b>	

Figure 4.2. Partial IPA consonant chart

By IPA convention, in paired voiced vs. voiceless sounds, the voiceless sound appears to the left of its voiced counterpart. Note that the blank box in the chart simply indicates that we do not deal with velar fricative articulations in this book. Many languages and language varieties have velar fricatives, like German, e.g. at the end of the word *Bach*, or Scottish English, e.g. at the end of the word *loch* (a word for ‘lake’, as in *Loch Ness*).

Different IPA labels referring to place of articulation can be combined to provide more detailed articulatory descriptions. For example, the sounds [f, v] are usually described as *labio-dental* fricatives, accounting for the fact that their pronunciation involves both the lips and the teeth.

Vowel space is usually plotted inside a diagram like the one in Figure 4.3, called a **vowel quadrilateral**.

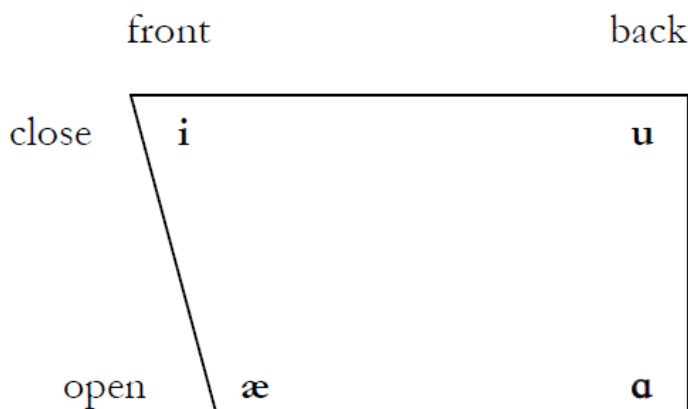


Figure 4.3. Partial IPA vowel chart.

The diagram represents the left profile of the inside of the mouth, the side that is also represented in standard diagrams of the vocal tract. Each of the four sides in the diagram represents the top, bottom, front and back of the oral cavity. The four vowels discussed in this chapter can be plotted at the four angles of the quadrilateral. The vowel quadrilateral accounts for tongue movement only. It does not represent lip movement.

### *DF representations of speech sounds*

Here is a summary of a sample of distinctive features.

		Yes	No
<b>Manner of articulation</b>	the airflow is smooth	[+sonorant]	[-sonorant]
	there is contact of articulators	[+stop]	[-stop]
	the air flows through the nasal cavities	[+nasal]	[-nasal]

		Yes	No
<b>Place of articulation</b>	the body of the tongue is raised	[+high]	[-high]
	the body of the tongue is lowered	[+low]	[-low]
	the body of the tongue is fronted	[+front]	[-front]
	the body of the tongue is retracted	[+back]	[-back]
	one or both lips are involved	[+labial]	[-labial]
	the teeth are involved	[+dental]	[-dental]
	the tongue blade/tip is involved	[+coronal]	[-coronal]
<b>Voicing</b>	the vocal cords vibrate	[+voice]	[-voice]

Figure 4.4. Sample DF labels

DF approaches typically use a matrix-like table to classify speech sounds, where vowels and consonants are plotted together.

The matrix in Figure 4.5 gives a classification of the speech sounds that are represented by the phonetic symbols in its first column. For example, the four vowel symbols [i, æ, ɪ, u] represent oral vowels, and the feature [nasal] is accordingly marked for these vowels with a minus symbol. Many languages have nasalised vowels (e.g. French, Portuguese, Hokkien), for which there are different phonetic

symbols from the ones given here, and whose [nasal] feature is therefore marked with a plus sign in a DF matrix. We marked the consonants [t, d, n] as [-dental] in the matrix, although they can be [+dental] in several languages and in several varieties of English. When you articulate [t, d, n] as [+dental], the tip of your tongue touches your upper teeth instead of the alveolar ridge. In all cases, [t, d, n] are [+coronal].

In the matrix below, we use the following conventions:

**son:** sonorant **lab:** labial **dent:** dental **cor:** coronal

	Manner of articulation			Place of articulation							Voicing
	son	stop	nasal	high	low	front	back	lab	dent	cor	voice
<b>p</b>	-	+	-	-	-	-	-	+	-	-	-
<b>b</b>	-	+	-	-	-	-	-	+	-	-	+
<b>t</b>	-	+	-	-	-	-	-	-	-	+	-
<b>d</b>	-	+	-	-	-	-	-	-	-	+	+
<b>k</b>	-	+	-	+	-	-	+	-	-	-	-
<b>g</b>	-	+	-	+	-	-	+	-	-	-	+
<b>m</b>	+	+	+	-	-	-	-	+	-	-	+
<b>n</b>	+	+	+	-	-	-	-	-	-	+	+
<b>ŋ</b>	+	+	+	+	-	-	+	-	-	-	+
<b>f</b>	-	-	-	-	-	-	-	+	+	-	-
<b>v</b>	-	-	-	-	-	-	-	+	+	-	+
<b>s</b>	-	-	-	-	-	-	-	-	-	+	-
<b>z</b>	-	-	-	-	-	-	-	-	-	+	+
<b>i</b>	+	-	-	+	-	+	-	-	-	-	+
<b>æ</b>	+	-	-	-	+	+	-	-	-	-	+
<b>a</b>	+	-	-	-	+	-	+	-	-	-	+
<b>u</b>	+	-	-	+	-	-	+	+	-	-	+

Figure 4.5. Sample DF matrix

It's worth highlighting that in both frameworks, IPA and DF, the different configurations of the articulators are independent of one another. For example, within the IPA framework, a plosive can be voiced or voiceless, labial, alveolar or velar. Similarly, within the DF framework, a [+high] sound can be [+front] or [+back], [+labial] or [-labial]. The important thing to keep in mind is that each movement

of each articulator produces one particular effect on the flow of air. It is the combination of *concerted* effects that produces a speech sound.

#### Activity 4.5

Using only the two DF [sonorant] and [stop], how would you characterise the set of sounds described by the following IPA labels:

plosives    nasals    fricatives    vowels

### *Redundancy in phonetic descriptions*

Our realisation that a sound is the result of a complex interplay of factors does not mean that its *description* must reflect this complexity. If description consistently matched real-world complexity, nobody would be able to learn about Newton's Laws of Motion in secondary school. Science of course strives to make sense of complex phenomena by means of **simple** descriptive statements. Let's illustrate this point with an example. From the DF matrix in Figure 4.5, we may conclude that the English sound represented by the symbol [n] can be described as:

(4.2) [+son +stop +nasal -high -low -front -back -lab -dent +cor +voice]

This description is accurate because it accounts for the uniqueness of [n] among all other sounds. But it is also, to say the least, cumbersome. It is like describing a cat, as opposed to a human being, as having fur, two eyes, four limbs, mammal features, claws,

a stomach, whiskers, feline features and small size. Some of these properties can be inferred from other properties. For example, “four limbs” can be predicted from “mammal features”. Other properties apply equally well to other members of a larger class. For example, both cats and humans have two eyes and a stomach. Likewise, in terms of distinctive features, [+nasal] implies [+sonorant], because nasal airflow is smooth (unless you’re blowing your nose!). Also, most sounds in Figure 5.5 share features like [-low -front]. An equally adequate description of the sound [n], in terms of distinctive features that apply to English is:

(4.3) [+nasal +cor]

This is so because, in English, there are only three nasal sounds, [m, n, ŋ]. Specifying their place of articulation is thus enough to describe the uniqueness of each nasal in this language. The description in (4.3) is as informative as the one in (4.2), and it is clearly simpler. We should stress, however, that (4.3) provides a description of [n] that applies to a single language, in this case English. In other languages, the set of nasal sounds may include vowels, as we noted above, or other consonants besides [+stop], or [+stop] nasals that are [+coronal] and [+dental]. The analysis of sounds in different languages must of course take into account the specific articulatory patterns that are found in each language.

Our discussion of the description in (4.2) highlights two important points. The first is that a DF analysis provides us with more labels than are necessary for the description of single sounds in particular languages. This is also true of the IPA approach. We saw that the combination of velar and fricative articulations does not exist in English, although velar and fricative are useful labels to describe other articulatory combinations in this language. In other words, both IPA and DF frameworks contain some degree of redundancy, in that they specify more structure than is required for the actual description of any given speech sound. A complete IPA and DF chart of the full set of articulations involved in English, for example, would show greater redundancy still. We will see another example of redundant analysis in section 6.3.1, where a theoretical “slot”

was posited for inflectional prefixes, a unit that is not needed for the analysis of English morphemes, but which occurs in other languages. Redundancy is a feature of any scientific analysis, and of taxonomies in particular. For example in zoology, mammal and eight legs are constructs that usefully describe different types of living organisms, although there are no eight-legged mammals. Redundancy is also a feature of language itself, in fact an invaluable one in everyday communication.

The second point that we wish to note is that redundancy operates at two different levels, the level of individual languages and the level of language in general. We saw above that (4.3) removes the redundancies found in (4.2) for the description of one sound in one language. We said that one of the reasons why we were able to simplify (4.2) is that some features are predictable from other features. In English, [+nasal] predicts [+stop], because all English nasals involve a complete closure in the mouth, blocking the airflow. But we also said that [+nasal] implies [+sonorant], because nasal airflow is smooth, and this is a universal entailment. It applies to nasal sounds in any language. An articulation like \*[+nasal -sonorant] is therefore not generally found: we already know that not all theoretically possible articulations occur in practice. This is often due to simple physical reasons, that apply to any human being and therefore to any language. For example, try producing a speech sound which involves contact between your soft palate and your lower lip!

#### Activity 4.6

1. Which of the following (sets of) features necessarily implies the other? Circle your answers.



The symbol  $\Rightarrow$  stands for ‘implies’.

(a) [+high] $\Rightarrow$ [-low]	(b) [-low] $\Rightarrow$ [+high]	(c) [+stop] $\Rightarrow$ [-son]
(d) [+voice] $\Rightarrow$ [+son]	(e) [+son +stop] $\Rightarrow$ [+nasal]	(f) [+cor] $\Rightarrow$ [-dent]

2. Which of the following combinations are impossible to articulate? Circle your answers.

(a) [+dental +coronal]	(b) [+front +back]	(c) [-front -back]
(d) [+sonorant +stop]	(e) [+nasal -stop]	(f) [+cor +lab]

#### Activity 4.7

In other languages that you speak, try to find features that entail one another are redundant for the description of particular sounds

The clear differences, that we have highlighted, between IPA and DF analyses should not obscure the fact that their object is the same: the articulation of human speech sounds. For example, the same sound [p] can be identified as a voiceless bilabial plosive or as [-son +stop -voice +lab]. These descriptions do not reflect a

difference in the articulation of the sound [p]. They reflect a preference for a set of labels and representations that each framework, or theory, finds more useful for the purposes of their investigation.

Despite the different assumptions and insights into articulation that each of the two framework offers, the phonetics literature does not always show a clear-cut choice between one or the other. As we noted at the beginning of this section, phoneticians may draw on terminology from either account, when describing their research. This is often so for the sake of simplicity. For example, it is quicker (and neater) to say plosive than to say [-son +stop]. But it is also simpler to say [+sonorant] than to say *vowels* and *nasals*. This terminological hesitation reflects the fact that in phonetics, as in any science, the search continues, for the most economical way of making sense of our observations.

#### Activity 4.8

Take some time to study the IPA and DF charts given in this chapter, and to provide minimal descriptions of each sound, in order to make clear to yourself what makes each sound unique. Make sure that you understand the reasons for each of the labels applied to each sound. Re-reading the Try the following sections in this chapter may help you with tips to carry out these tasks.

## 4.5 Intonation and tone

Vowels are voiced sounds, produced with a smooth airflow. These

two features make vowels the sounds upon which intonation is chiefly modulated.

Intonational modulation, or speech melody, is one crucial carrier of linguistic meaning in any language.

Try the following

Imagine yourself in two different situations where you might use the following utterance, one jokingly, one deeply annoyed:

*Oh, shut up!*

If you try to pronounce this utterance to match the suggested feelings, you'll find that you pronounce the "same" utterance in two very different ways. In fact, you'll be pronouncing two different utterances, with two distinct meanings.

**Intonation** can be said to result from rapid changes in the tension and rate of vibration of the vocal cords. The vocal cords are folds of elastic tissue that can stretch or contract, and vibrate at a higher or lower rate, much like violin strings. Associated with these changes, the length of the vocal tract may be modified during speech, due to the vertical mobility of the larynx.

Try the following

Touch your larynx gently with your fingers, or look in a mirror with your chin lifted up, so you can see your larynx clearly. Now sing the highest note you can sing, and then the lowest, and then go from one to the other, a few times in a row. You'll feel/see your larynx moving up for the high note, and down for the low one, assisting the higher and lower rate of vibration of the vocal folds, respectively. The effect that these combined actions have on the column of air inside your vocal tract gives a similar impression to the one achieved by a "bottle concert", where you fill similar bottles with different amounts of liquid and then either blow into them

or strike them with some hard instrument. You'll get a higher note when there's a higher rate of vibration inside the bottle and inside your vocal tract, and lower notes for lower rates.

Like any other linguistic system, intonation systems vary both across and within languages. However, two basic distinctions in the form of intonation patterns appear to be consistently associated with two sets of meanings, according to whether our tone of voice falls or rises.

### *Falling tones*

**Falling tones**, or *falls*, result from a decrease in the rate of vibration of the vocal cords. Falls are typically associated with statements and commands. These types of utterances require minimal or no verbal response from the listener, and falling tones are therefore said to convey a **closed** set of meanings, in that they close off communication.

### *Rising tones*

**Rising tones**, or *rises*, result from an increase in the rate of vibration of the vocal cords. Rises typically signal questions, which call for some response from the listener, and continuation, when the speaker goes on speaking. That is, the use of rising tones in an utterance indicates that the utterance is not complete, and rises are therefore associated with an **open** set of meanings.

#### Activity 4.9

Try pronouncing a single word, e.g. Yes, with as many different intonation patterns as you can. Check for yourself or with a friend the range of different meanings that the “same” word can convey. Here are a few suggestions: say the word with a falling tone, then a rising tone; say the word with a very long fall, followed by a very long rise; say the word with a fall-rise pattern; say the word with a rise-fall pattern.

Or try “reading” out the following versions of Yes:

Yes! Yes? Yes!!!Yes??Yes?! Yes!!!!? Yeeees!

If you have difficulty producing (and hearing) falling vs. rising tones, or level tones, try Collins and Mees’ (2003: 118) analogy: “The engine of a motor car when ‘revving up’ to start produces a series of *rising* pitches. When the car is cruising on the open road, the engine pitch is more or less *level*. On coming to a halt, the engine stops with a rapid *fall* in pitch.”

Intonational modulation operates across whole utterances, including those that may comprise a single word. In many languages, the difference between a statement and a question is signalled by intonation alone, a fall vs. a rise at the end of the utterance, respectively, without the associated change in word order that is common in standard uses of English.

All languages make intonational distinctions of this kind, but there are other uses of pitch in language. Pitch modulation may operate on single words, distinguishing between different lexical meanings of the same string of vowels and consonants. This is the domain of **tone**. Here is one example from Ngbaka, spoken in the Central

African Republic. This language has level tones, meaning that the tone is kept even, neither rising nor falling, throughout the pronunciation of the words:

**Ngbaka**

Word	Tone	Meaning
[ma]	low, level	‘magic’
[ma]	mid, level	‘I’
[ma]	high, level	‘to me’

Figure 4.6. Examples of tones in Ngbaka

Asian languages like Mandarin, African languages like Hausa, and several American Indian languages are **tone languages**, distinct from **intonation languages** like English or Malay. Other languages, like Swedish and Japanese, combine features of intonation and tone languages.

As a concluding thought, it is interesting to note that despite the crucial role played by pitch modulation across all languages, the meanings conveyed by such modulation are hardly represented in spelling. Question marks vs. exclamation marks, for example, give an even rougher approximation to the various rises and falls of real-life language than the English spellings *bus*, *cress*, *science*, *city*, *mouse*, *flaccid*, give to the sound [s]. The deficient representation of intonation in print highlights the very limited resources of written forms of language to convey the full range of linguistic meanings. It is often the case that the meaning that you intend for an utterance is in fact given by intonation, rather than by the words that you use. We’ve all heard the saying: it’s not *what* you say but *how* you say it that counts. Or, as the American linguist Kenneth Pike (1945: 22) strikingly put it, “[...] if a man’s tone of voice belies his words, we

immediately assume that the intonation more faithfully reflects his true linguistic intentions.”

#### Activity 4.10

Can you say this utterance with two different intonations, so that it conveys sincerity in one case, and irony in the other? Now, that was clever!

#### Activity 4.11

Can you say this utterance so that two different meanings are possible, one where the speaker did in fact go out, and the other where the speaker in fact stayed in? I didn't go out because I wanted to see you.

In this chapter, we have provided a very brief overview of the range and variety of speech sounds that can be produced by the human vocal tract. Obviously, no language makes use of all possible speech sounds. In the next chapter, we look at how speech sounds are organised and used in languages.

## Food for thought

When it's English that we SPEAK

Why is STEAK not rhymed with WEAK? And couldn't you please tell me HOW COW and NOW can rhyme with BOUGH?

I simply can't imagine WHY HIGH and EYE sound like BUY.

We have FOOD and BLOOD and WOOD, And yet we rhyme SHOULD and GOOD.

BEAD is different from HEAD,

But we say RED, BREAD, and SAID. GONE will never rhyme with ONE

Nor HOME and DOME with SOME and COME.

NOSE and LOSE look much alike,

So why not FIGHT, and HEIGHT, and BITE?

DOVE and DOVE look quite the same,

But not at all like RAIN, REIN, and REIGN.

SHOE just doesn't sound like TOE, And all for reasons I don't KNOW. For all these words just prove to ME That sounds and letters DISAGREE.

## Further reading

Collins, Beverley and Mees, Inger M. (2003). How we produce speech. In *practical phonetics and phonology: A resource book for students*. London/New York: Routledge, pp. 25-39.

(This book includes a CD with sound files to listen to and practise with.) Deterding, David H. and Poedjosoedarmo, Gloria R. (1998). Chapter 2.

Speech production. In *The sounds of English. Phonetics and phonology for*

*English teachers in Southeast Asia*. Singapore: Prentice Hall, pp. 9-13.



Roach, Peter (1993). Chapter 2. The production of speech sounds. In *English phonetics and phonology. A practical course*. Cambridge: Cambridge University Press, pp. 8-17.

## References

Collins, Beverley and Mees, Inger M. (2003). How we produce speech. In *practical phonetics and phonology: A resource book for students*. London/New York: Routledge.

Pike, Kenneth L. (1945). *The intonation of American English*. Ann Arbor: University of Michigan Press.

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.

# Chapter 5: Morphology - The grammar of words

## The grammar of words: words and word parts

### Chapter Preview

What is a word?

Why are there different types of words?

Can words contain other words?

Can words contain other meaningful elements that are not words?

## 5.1 Introduction

In this chapter, we discuss how linguistic meaning is encapsulated in different levels of patterning of linguistic form, looking in turn at words, sounds and sentences. We will do this by looking at language

data that illustrate the particular points about linguistic meanings that we wish to highlight. Much of the data will be taken from English, because this is the language that we share: we are writing this book in English and you are reading it in English. Do keep in mind, however, that English is one particular language among thousands of other languages around the world, and that this book deals with how to analyse language in general, not particular languages. This means two things.

First, the language data used in this book are to be taken as **representative** examples. Clearly, no one can use an infinite amount of data in any book. We have chosen the best examples we could think of, to help us bring our points across in the clearest possible way. Here's an example of what we mean. If you had to explain to us what an insect is, you could choose to show us samples, like a beetle or an ant, as representative members of a class of beings called "insects", and then tell us to work out, from these samples, what is it that characterises insects. You could even add that a spider or an earthworm are not insects, to help us better understand what can rightfully be called an insect and what cannot. This is exactly what we expect you to do from the examples that we give.

Second, the concepts that we exemplify using the English language are meant to apply to other languages as well. Think of our English examples as you would think of using the example of a pendulum, or of a coin dropping to the ground, to explain the general law of gravity. It wouldn't make sense to conclude that the law of gravity applies only to coins or to pendulums, just because we use coins and pendulums to demonstrate the effect of the law.

One very important way in which we human beings understand the world around us and learn to talk about it is by means of **generalisation**. We find certain properties in certain objects, and we extend those properties to other objects that we perceive as similar. This is why some people say that "everybody" knows that

“women” enjoy shopping, or that “snakes” are dangerous. No one has ever polled every single woman to check their shopping enjoyment, or every single snake around the world to check their threatening behaviour. Nor every single human being about their opinions on women or snakes. Rather, generalisations such as these are based on observation of a subset of the population of women and snakes, respectively. Here’s another example to bring home the same point. If you taste cheese for the first time and find it disgusting, you may conclude that any kind of cheese will taste disgusting. Generalisation is a vital component of any scientific explanation, but with one proviso, that is equally vital. Just as you would have to change your mind about cheese if you decide to try a Camembert and find it delicious, so scientists will narrow down their generalisations, and change the labels by which they call things, when they find **counterexamples**, examples that contradict what they previously thought. In the same way, what we will have to say in this book about English applies generally, across different languages. A good way of reminding yourself of this throughout this book is to try to apply the concepts that we describe here to any other language(s) that you may be familiar with, as soon as they are described. At the same time, you should also be looking out for counterexamples, whether in English or in other languages that you are familiar with. A few activities and exercises in the book will help you keep these two very important points in mind!

## 5.2 The word “word”

**Morphology** can be defined as the study of words. Let’s start by checking out what a word is.

(5.1)

A very long (Welsh) word

*Llanfairpwllgwyngyllgogerychwyrndrobwlllandysiliogogoc  
h*

A medium-sized (multilingual) word

*karaoke-singers*

A very short (English) word

*a*

You may or may not agree that all three examples above are words. If you feel any discomfort about calling each of them a “word”, you are not alone. **Word** is in fact one of the concepts in linguistics that defies precise definition. Since everyone talks about words, we might assume that this is a well-understood concept. The truth is that it is not. What is meant by a word varies greatly from language to language. Some languages, like Mandarin, have largely monosyllabic words, which are words that consist of a single syllable. Other languages, like Malay, allow the stringing together of several syllables, or several words, to form larger words. This is also the case in Welsh, where the word in (5.1) contains other words, *llan*, *fair*, *pwll*, etc., just like many words of English do too. For example, the English word *handbag* contains the word *hand* and the word *bag*. The difference is that Welsh allows the spelling together of more words than English does. If English had the same spelling rules as Welsh, the English translation of the Welsh word above could be spelt like this:

*Churchofsaintmaryinthehollowofthewhitehazeltreesnearthefi  
ercewhirlpoolandthe churchofsainttysiliobyaredcave.*

Similarly, the Swedish word *ettusentrehundrasjuttiofyra* could be spelt:

*onethousandthreehundredandseventyfour*

Trying to work out the meaning of a word so spelt is not much worse than trying to work out the meaning of chemical compounds or of medical terms, like this well-formed English word that we found listed in the Oxford English Dictionary:

*pneumonoultramicroscopicsilicovolcanoconiosis*

What is meant by a “word” also varies within a single language, including whether we are talking about speaking words or writing them. Perhaps you would want to say that *karaoke-singers* is two

words, because of the hyphen separating *karaoke* from *singers* in the spelling? Or would it satisfy you better to say that it is a single word, because it has a unique meaning, just like say, *sopranos* does? This is a single word whose meaning can be rephrased as ‘opera-singers’. The best definitions of the word *word* take either orthography (spelling) or rhythm into account. Orthographic definitions take *word* as a unit that is separated by a blank space on each side, in a printed text. Rhythmical definitions take it as a unit of speech that can be separated by an optional pause, meaning that a word can be pronounced on its own, preceded and followed by silence. The blank spaces on a page represent the possible silences between words. Robot-characters in some science-fiction films pronounce sentences in just this way, word by word.

#### Activity 5.1

Try to pronounce the following sentences in a robot-like manner, word by word: I saw a black bird by the green house. I saw a blackbird by the greenhouse. How many words do you count in each sentence?

The fact that we do understand what robot-like sentences mean, despite their stilted delivery, reveals one very important feature of human speech. This is that speech is a stream of ordered units, as already discussed in Chapter 1. Human speakers, and robots that attempt to replicate their speech, must order the units that make up their utterances in a particular way, so that the utterances so formed make sense. That is, linguistic units must occur in particular positions along the stream of speech, surrounded by other units. This observation helps us clarify one concept that is central to linguistic analysis, the concept of distribution.

## 5.3 Distribution

Linguistic units, like people or objects, show up in predictable places, or contexts. You wouldn't expect to find the morning newspaper tucked away in the fridge, or a cat sitting next to you watching a movie at the cinema, or a palm tree in full bloom in the North Pole. Linguistic units pattern in the same predictable way.

The **context** of a particular linguistic unit is given by the linguistic units of the same type that surround that unit. For example, the context of a particular word is given by the words that precede and follow it. The same holds for morphemes (see section 5.4.2 below for *morphemes*) or sounds. Analysis of context in these terms is a natural consequence of the fact that speech occurs along the dimension of time: sounds follow sounds, words follow words. The **distribution** of a particular linguistic unit is then the set of contexts in which that unit is found to occur. By the same token, you can also work out your own distribution, if you list all the places in which you are likely to be found.

The distribution of linguistic units may be represented by a **distributional frame**. Given a form like, say, XYZ, the context of the unit Y is given by a distributional frame with the general form:

X \_Z

This representation is similar to a more familiar one like, say,  $a + b = c$ , where each letter stands for a number. In both cases, the letters represent a variable that can be replaced by something else, according to conventions that all users of these representations have agreed upon. The conventions in distributional frames are:

- Each symbol X, Y, Z, represents one linguistic unit of the same type, e.g. a morpheme or a sound. These units occur in the given sequential order: X precedes Y, and both precede Z.
- The blank represented by indicates the context in which Y, the unit in question, occurs. In this case, Y follows X and precedes

Z.

For example, given the phrase *the brown cat*, we say that the distribution of the unit (word) *brown* is given by the distributional frame *the \_ cat*, where the blank indicates the occurrence of *brown*. Likewise, the distribution of the *in* the same phrase is given by the frame *\_ brown*. Or, given the word *cat*, pronounced [kæt], we say that the distribution of the unit (sound) [æ] is given by the distributional frame [k] \_ [t] (see Chapter 3 for conventions on the representation of speech sounds).

## 5.4 Morphological units

There are two chief levels of word patterning that interest morphologists. One deals with patterns of words to form phrases or sentences, based on our observation that different words behave differently when chained together with other words. The characteristic behaviour of particular groups of words allows us to classify words into different word classes. The other level of morphological analysis deals with patterns within the words themselves, that is, with the internal grammatical structure of words.

We now discuss these two analytical levels in turn. Notice, however, that both analytical levels work together to provide us with insight about word patterning.

### 5.4.1 Word classes

Check the following data:

(5.2)



*The cat sleeps on the mat.*

*\*The sleeps cat on the mat.*

*\*The cat sleeps the on mat.*

Recall, as said in Chapter 1, that an asterisk preceding a form indicates that that form does not occur in the language under analysis. All three sentences in (5.2) contain the same words. The only difference is that the words pattern differently, resulting in one well-formed sentence and two ill-formed ones. For example, we see that the word *the* can precede the word *cat*, but not the words *sleeps* or *on*. Observations of this kind lead us to assign words to different **word classes**, according to their relative positions in a phrase or a sentence, that is, according to which words can or cannot follow or precede other words. This chain ordering of words tells us about their **distributional properties**, by showing us which particular positions are grammatical for which types of words. In the linguistics literature, *word classes* are sometimes called *parts of speech*, *lexical classes*, *grammatical categories*, *grammatical classes*. For our purposes, you can take all of these terms as equivalent.

Words can be broadly divided into two main word types – **lexical words** and **grammatical words** – according to their distribution and their meaning. Within each of the two types, several word classes can be further distinguished, because each word class patterns in characteristic ways.

Below, we give a number of principles, or criteria, that can help us identify different word classes. As you read through them and think about their application to different words in the same word class, you should keep in mind that none of the current criteria for defining word classes is watertight, including the ones that we suggest here. This means that once you understand how these criteria apply, you will be able to come up both with words that fit the criteria as well as with words that fail the criteria. The latter are **counterexamples**, showing that the criteria represent generalisations about what holds true most of the time, rather than

all of the time. We will give a few examples of these ‘bad-behaved’ words ourselves, where relevant. This is not a problem for our analysis: this book presents only a very elementary set of principles to help us deal with language (languages are very, very complex things!), and professional linguists themselves go on being baffled by the complexity and quiriness of language. One of them, Edward Sapir, once famously said that “all grammars leak” (Sapir 1921: 38). These “leaks” are precisely what makes us want to go on trying to understand how language works, so that we can fix them to make our analyses “flow” in a satisfactory way. In this sense, a good grammarian is like a good plumber. Take the criteria that we offer here as *typical* criteria, that do useful work in helping us identify word classes, but only in the majority of cases rather than in all cases.

### *Lexical word classes*

Lexical words represent a specific **referent** in the world of our experiences. They refer to objects and substances that we can see, sensations that we can feel, qualities and events that we can observe. Lexical words therefore form the largest group of words in languages. The word classes to which they belong are **open classes**, in that the overwhelming majority of new words that become part of a language are of this type. In the linguistics literature, the terms *lexical words*, *content words* and *open-class words* are sometimes used interchangeably. These terms reflect the fact that whenever a new interesting thing is created, invented or found, we immediately create, invent or find a new lexical word to go with it, so that we can talk about it. When the authors of this book were growing up, there was no email, SMS or blogs, and so there were no words for these things either. There were, however, telegrams, telex and vinyl LPs, and the words for them were part of daily life at the time. New words that restock the vocabulary of languages, and old words that fade away for lack of linguistic demand are all lexical words. We can

say that languages keep themselves alive and working through the comings and goings of their lexical words.

Lexical meanings are generally referential and, as such, they are arbitrary and idiosyncratic. Four lexical word classes can be distinguished in English according to their distributional properties. These are nouns (N), verbs (V), adjectives (Adj) and adverbs (Adv). We now discuss each one in turn.

## Noun (N)

Nouns are the only lexical word class that can be followed by a mark of plural, in English. Plural forms of nouns are often represented in spelling by – (e)s at the end of the word in so-called regular plurals, where (e) may or may not occur, depending on the spelling of the singular form. For example, the word *cup* is a noun in the singular form, representing a single object that we call “cup”. The word *cups* is also a noun, in the plural form that indicates more than one of those objects. The same is true of a noun like *brush* and its plural *brushes*. Regular forms are forms that are productive, or active, in languages. Taking English nouns as example, this means that whenever a new noun comes into the language, its plural will be formed by adding –(e)s to it, as in *faxes*, *emails* or *modems*.

### Activity 5.2

*Which mouse?*

The plural of mouse is *mice*, right?

But what is the plural of *computer mouse*??

By our distributional criterion, if you can fit a word in the following frame, then that word is a noun:

*One , several (-e)s*

The regular plural marker is given in brackets to allow for irregular plural nouns like *foot-feet*, *tooth-teeth*, and *goose-geese*, as well as nouns which have the same form for singular and plural, e.g. *sheep*, *deer*, *fish*, and *fruit*. We do not otherwise deal with irregular morphology in this book.

Our distributional frame shows that a word like *cow* (or *foot*) is a noun. However, we need to point out that saying that “only nouns can be pluralised” does not mean that ‘all nouns can be pluralised’. English nouns pattern in two different ways. Nouns like *cow* can be pluralised because their referents can be counted. These nouns are therefore called **count nouns**. A word like *milk* cannot fit the given frame, because phrases like *\*one milk* and *\*several milks* are not well-formed for most speakers of English. But words like *milk* fit other distributional frames associated with nouns, e.g. patterning after a determiner (see below). The word *milk* is therefore a noun too, although its referent cannot be counted. The reason is that nouns like *milk* refer to shapeless substances, and counting applies only to referents with well-defined dimensions. This being so, we can propose a frame that provides a shape- giving context for non-countable nouns:

A \_ of \_

The first blank can be filled with countable, shape-giving words like *box*, *packet*, *loaf*, *glass*, *bowl* (sometimes called *measure words*), much in the way that classifiers are used before nouns in Chinese and other Asian languages. The non-countable noun fills the second

blank. For example, *a glass of milk* is a well-formed English phrase. “Shapeless” nouns like *milk* are called **mass nouns**.

### Activity 5.3

Can we use the following frames to help us distinguish count nouns from mass nouns?

many \_\_\_\_\_ much \_\_\_\_\_

Which of these frames would you choose for nouns like luggage, cattle, sugar, coffee, tea?

Besides plural forms, another feature that characterises nouns is that they can be followed, in writing, by -’s, a mark that is used to indicate possession, or belonging, and that is called *possessive* or *genitive* in the literature. When used with a noun followed by another noun, the possessive mark indicates that the second noun in some way is part of, or belongs to, the first one. For example, when we refer to *the cat’s tail* we are talking about the tail of a particular cat.

Nouns name entities of various kinds. Nouns that name people, places, institutions and brands, like *Jane*, *Malaysia*, *Telecom* or *Stradivarius*, are called **proper nouns** (or *proper names*) and are also treated as nouns in the literature. All other nouns are **common nouns**. Like common nouns, proper nouns can be followed by the possessive ’s. One example is a phrase like *Malaysia’s climate*.

However, the remaining properties of proper nouns are different from those of common nouns. First, proper nouns cannot be said to

have well- established referents across the board. The name *India*, for example, refers to a particular country, but the word *India* does not “mean” the country named by this word, nor does it mean a set of objects with perceivable shared properties. Rather, it is a label that the country goes by, just like *Mary* is a label that people called *Mary* go by. It would be hard to find a common feature of meaning among all people named *Mary*, parallel to the feature of meaning that allows us to designate all cups by the name “cup”. What all the individuals named *Mary* have in common is that someone decided to call them “Mary”.

Second, these words fit only marginally in both of the distributional frames given above. They do not designate substances, and we can “count” *Jimmies* and *Chloes* in particular groups (children love to do this), or *Ugandas* and *Englands* if we want to highlight, say, striking features of different parts of the same country (adults love to do this), but only in a marginal sense.

Third, proper nouns pattern equally marginally with other words that can precede nouns, like determiners and adjectives (see below): expressions like *That Japan is my favourite one* or *This is a green Matthew* are, to say the least, unusual. Given these provisos, we follow here the traditional classification of proper nouns as nouns.

## Verb (V)

Verbs can be followed by a mark of past tense, often represented in spelling by *-(e)d* in so-called regular verbs. As noted above, we are only concerned with regular morphology. A word like *bake* fits this frame, and is therefore identified as a verb:

*Today I / it , yesterday I / it (-ed)*

Some sentences contain only one verb. Examples are *I tripped* or *My neighbours have two dogs*, where *tripped* and *have* are the verbs. Other sentences contain several verbs, from two up to a maximum of five, in English. Examples are *Janet is singing* and *The laundry would have been being washed*, where *is singing* and *would have been*

*being washed* are all verbs. Some people doubt the “correctness” of the latter string of verbs, but there’s nothing grammatically wrong with it: it is simply a less common construction.

We call the last verb in strings like these the **main verb**, because this is the verb that carries the referential meaning that is being talked about. The other verbs are called **auxiliary verbs**, or **auxiliaries**, because their function is to help specify the time and duration of the action or state indicated by the main verb. In sentences with only one verb, this verb is of course the main verb. For example, the sentence *Janet is beautiful* has only one verb, *is*, and this is the main verb.

Three English verbs, *be*, *do* and *have*, can function either as main or auxiliary verbs. In the sentence *My neighbours have two dogs* the verb *have* is the main verb, whereas it is an auxiliary in the sentence *The laundry would have been being washed*. Because of this dual function, these verbs are very common in English, and because they are so common they are also very irregular. We saw above that newcomer words to a language, like *email*, that are therefore just becoming common in it, follow regular patterns, not irregular ones. Words are like clothes, the more you use them the more shapeless they get and the less they look like new clothes off the rack. The verb *be* is in fact the most irregular verb in English, in that it can appear in eight different forms. Just to satisfy your curiosity, these forms are: *am*, *are*, *is*, *was*, *were*, *being*, *been* and *be* itself!

Verbal forms that vary according to tense (and/or person, see below) are called **finite** forms. For example, forms like *are*, *were*, *baked*, *bakes*, are finite. If a sentence contains only one verb, the form of that verb is finite. The verb *have* is finite in the sentence *My neighbours have two dogs*, because it changes to *had* when we talk about the past, or to *has* when we say *My neighbour has two dogs*. Verbal forms that remain unchanged regardless of tense and person are **non-finite** forms. For example, *(to) speak*, *speaking*, *spoken*, and *(to) be*, *being*, *been*. In a sentence with main verb and auxiliaries, the first auxiliary is finite and all the other verb forms are non-finite.

In the sentence *The laundry would have been being washed*, only the auxiliary *would* is finite.

#### Activity 5.4

Indicate whether the underlined verbs are used as main verb or auxiliary.

1. I am doing my homework.
2. I am a teacher.
3. I have been doing my homework.
4. I have seen that film.
5. I did my shopping yesterday.

Now indicate whether the same verbs are finite or non-finite.

#### Activity 5.5

Can you explain the language play in this dialogue?

Speaker A. *Time flies!*

Speaker B. *I can't, they fly too fast!*

Hint: the play has to do with nouns and verbs.

### Adjective (Adj)

Adjectives pattern in two typical ways. They can immediately



precede nouns, or they can follow forms of a verb like *be*. The distributional frames for these two patterns are:

A \_\_\_\_\_ cow                      This cow is \_\_\_\_\_

If a word can pattern in these two ways, then it is an adjective: *brown* is one example. Typical adjectives do indeed pattern in these two alternative ways, although a few do not. Some adjectives can be used only before nouns. For example, you can say *the current president*, where *current* is an adjective, but you cannot say *\*the president is current*. Conversely, other adjectives pattern only after *be*-like verbs: you can say *dinner is ready*, but not *\*a ready dinner*.

The function of adjectives is to modify nouns. This means that adjectives tell us something about a noun, specifying a property or a quality of that noun. If we say *the green book* or *that large whale*, we're indicating which particular book or whale we mean: green, not red, and large, not tiny. This qualification of nouns can also be a matter of **degree**. If we see first a large whale, and later a very, very large whale, we can express this by saying that the second whale is *larger* than the first. If we then see a third, absolutely enormous whale, we can say that this one is the *largest* of them all. In both cases we are comparing sizes, and grading them: one is larger than another, like one can be smaller than another, or one can have the greatest (or smallest) size compared to all the others. The technical names for these uses of adjectives are **comparative** and **superlative**, respectively.

Degree forms of adjectives vary according to the length of the adjective word. There are two rules. For short adjectives with one or two syllables, we add *-(e)r* and *-(e)st* at the end of the word. For long adjectives, we add separate words, *more* and *most* before the adjective. We can say that *dolphins are more intelligent than whales*, or that *they are the most intelligent marine mammals*, but we don't say that they are *\*intelligenter* than whales, or the *\*intelligentest* of all. By the same token, we say *easier*, not *\*more easy*, and *easiest*, not *\*most easy*. (And yes, this short word vs. long word rule has several exceptions: short adjectives like *tired* and *ready* are examples.)

## Adverb (Adv)

Traditionally, the “class” of adverbs is a sort of ragbag: if morphological criteria cannot clearly identify the class of a lexical word, the solution is often to call it an “adverb”.

This is not simply an easy way out of a difficulty in classifying certain words. It underscores the difficulty itself instead. Similar problems of classification usually arise for two main reasons: either we are basing our classification framework on insufficient data, or the classification framework itself needs revamping. In other words, either the number of clearly identifiable words that we have observed, large though it may be, is not enough to help us decide the word class of a particular word; or perhaps we should start thinking about discarding the class “adverb” altogether, and create new word classes that better explain what makes “adverbs” special. Zoologists, for example, had to create a new zoological class to account for platypuses, the very odd mammals that lay eggs instead of giving birth to live young as “regular” mammals do. Facts, about living beings or language, cannot be changed to suit a theory that doesn’t explain them. It is the theory, no matter how respectable or how popular, that must change in order to serve our understanding of these facts. Adverbs may well be the platypuses of the grammatical zoo.

The facts are that attempting to identify adverbs on formal or distributional grounds is not straightforward, in that they may vary widely in shape as well as in patterning within phrases and sentences. The most common criteria used in the definition of adverbs are of two kinds. One, a referential criterion, states that the meaning of adverbs can modify the meaning of different words or phrases, including other adverbs. That is, their meaning attributes some quality to the meaning of another word. For example:

(5.3)

*She is very pretty.*

(5.4)

*She speaks very fast.*

In (5.3), *very* modifies the adjective *pretty* by intensifying its meaning. In (5.4), the word *fast* modifies the verb *speaks* by specifying the way in which the speaking takes place, and *very* in turn modifies *fast*. Both *very* and *fast* are therefore adverbs.

The second criterion is distributional, and states that adverbs are generally mobile words. This means that they may occur in different positions within an utterance without loss of grammaticality. For example:

(5.5)

*Sadly, she is an idiot.      She, sadly, is an idiot.*

*She is, sadly, an idiot.      She is an idiot, sadly.*

The mobility of *sadly* identifies it as an adverb. Note that mobility is further demarcated by pauses, in speech, and by commas, in print. Note also that this criterion would fail to identify *very* as an adverb.

#### Activity 5.6

What is the adverb *sadly* modifying in this sentence?  
*Sadly, she is an idiot.*

### *Grammatical word classes*

We said above that lexical word classes express content in terms of concepts of various kinds. In contrast, grammatical words represent distributional relationships between lexical words. Grammatical words belong in **closed classes**, in that new words of a language rarely are of this type. Across languages, the number of words in grammatical word classes is therefore much smaller than in lexical word classes. In the literature, other terms that designate grammatical words are *closed-class words* and *function words*.

The meanings of grammatical words can be said to be structural and systematic, rather than referential and idiosyncratic. These meanings are the “glue”, as it were, that helps us express and understand the meaning relationships between lexical words. Four classes of grammatical words are usually described for English. These are determiners (Det), pronouns (P), conjunctions (C) and prepositions (P), as follows.

### Determiner (Det)

Determiners can precede both N and Adj: *this cow* and *this large cow* are both well-formed in English, where *this* is a determiner. A frame for Det might then be, where the brackets indicate an optional word:

(Adj) N

This frame identifies words like *the*, *a*, *these*, *my*, *your*, *both*, as determiners.

#### Activity 5.7

The word class determiner includes different types of words that fit the frame above. Would the distribution of determiners in the following data allow us to say that there are different types of determiner in English? Why?

Note: all the words in the examples are determiners, except friends and cars.

all my friends	*my all friends	
my three cars	*three my cars	
all my three cars	*my all three cars	*all three my cars
all three cars	*three all cars	

### Pronoun (Pr)

Pronouns can replace sequences of an optional determiner, followed by an optional adjective, followed by a noun. Using the symbol  $\approx$  to indicate distributional equivalence, we can formalise the distribution of pronouns in this way:

(Det) (Adj) N  $\approx$  Pr

In the three sentences in (3.6), the word sequences in italics can all be replaced by a word like *they*, identifying this word as a pronoun:

(5.6)

*Those black cats* are so annoying.

*Black cats* are so annoying.

*Cats* are so annoying.

Pronouns are interesting words, because they work by proxy, as it were. Hence their name, *pro*-nouns. **Personal pronouns**, for example, are used to replace direct reference to ourselves and our interlocutors. Instead of referring to ourselves and our conversation partners by name, we use pronouns. The pronoun *I*, for example, is used by speakers to refer to themselves, while the same speakers use *you* to refer to their listeners. When listeners in turn become speakers, they use the two words in exactly the same way. These are therefore called **first person** and **second person** pronouns, respectively – the speaker comes “first”. English has first person pronouns that have singular and plural forms, *I* to refer to the speaker only, and *we* to refer to a group including the speaker. In contrast, the English second person pronoun *you* has the same form to refer to one or more than one interlocutor. **Third person**

pronouns refer to what the conversation is about. In English, these pronouns have three forms for the singular, *he*, *she* and *it*, the first two usually referring to sexed beings, and the latter to inanimate referents or beings whose sex is irrelevant or unknown. The plural form is the same for all three, *they*.

Other pronouns are also used in relation to the participants in an exchange. **Possessive pronouns** (e.g. *his*, *hers*, *mine*, *theirs*, *yours*) indicate whether something belongs to or is a characteristic of those participants. **Demonstrative pronouns** (e.g. *this*, *those*) show the distance between the speaker and the referent that is being talked about. Take, for example, the sentence:

(5.7)

*These are yours.*

This sentence contains two pronouns, *these* and *yours*. The first is demonstrative, and signals a referent that is near the speaker. (In contrast, the pronoun *those* signals a referent further away from the speaker.) The second pronoun is possessive, indicating that the listener owns whatever the speaker is referring to.

### Activity 5.8

What is the word class of the underlined words in these sentences? Explain how you reached your decisions.

1. That man really loves boiled squid.
2. I can't understand that.
3. My plate of barbecued squid is much tastier than his.
4. I find his choice of food very funny.

## Conjunction (Conj)

Conjunctions are linking words, “conjoining” other words or phrases in order to enable multiple occurrences of the same word class or

phrase. Conjunctions link units that are of the same type, e.g. an adjective with an adjective or a pronoun with a pronoun (section 7.5 deals with conjunctions in greater detail). In these two examples, the italicised conjunction *and* joins two pronouns in the first sentence, and two sequences of Adv Adj in the second sentence:

(5.8)

You *and* I need to talk seriously.

My cat is very beautiful *and* very stupid.

## Preposition (P)

Prepositions are linking words that typically show relationships of space and time between other words or phrases. They are followed by (Det) (Adj) N sequences, or by pronouns, which replace these sequences. The distributional frames for a preposition are:

(Det) (Adj) N Pr

These frames identify words like *in*, *under*, *despite*, *through*, as prepositions. The meanings expressed by prepositions can be referential, like those of lexical words. For example, the word *under* regularly means a location tucked below something on a higher level in space. Or a preposition like *during* typically refers to a period of time. Other prepositional meanings are not fully lexical. For example, a word like *in* does not mean anything that can be usefully generalised from uses like *in the house*, *in a moment*, *in conclusion*, *in fact*, *in neat rows*, *in Swahili*. Similarly to the problems raised in the classification of adverbs, discussed above, these observations about prepositional meanings raise other problems for our analytical framework. In this case, we may question the adequacy of a watertight distinction between “lexical” and “grammatical” word meanings, that leaves prepositional meanings scattered between both.

Once particular words have been safely assigned a word class, we can use those words as shortcut tests of the class of other words. This is the **substitution** criterion. By this criterion, any word that

can replace the noun *cow* in contexts where *cow* is found is also a noun. The same principle applies to the other word classes.

Here is a summary of the word classes discussed in this chapter.

	Word class	Examples
Lexical words	N	ant, flower, beauty, contradiction
	V	do, think, snore, drink, explain
	Adj	blue, tall, necessary, expensive
	Adv	really, quite, wonderfully, never
Grammatical words	Det	a, some, many, our, those
	Pr	she, mine, theirs, this, those
	Conj	and, as soon as, because, if, however
	P	in, on, into, from, in front of

Figure 5.1. Summary of word classes

And here is one example of a sentence containing all eight word classes:

Figure 5.2. Example sentence with eight word classes

er	H at	lo oks	ni ce,	b ut it	alw ays	sle eps	o n	th eir	fin est couch
et	D N	V	A dj	C onj Pr	Adv	V	P	D et	Adj N



### Activity 5.9

Create a table consolidating the criteria that identify a word as belonging to each of the word classes above. Try creating new distributional frames for each word class.

### Activity 5.10

1. Choose any printed text (newspaper or magazine article, book, online article, etc.) and try to classify the words in one of its paragraphs into the word classes introduced in this chapter. Discuss any problems with a partner, or in a group.
2. Count the total number of lexical and grammatical words in your text.
3. Then count the number of repeated lexical words and repeated grammatical words that you found. How do your findings help you make sense of the labels open class and closed class, respectively, for these words?

## *Morphemes*

We saw above that words like *handbag* and *karaoke-singers* may contain words, each of which is meaningful in its own right. We also find other words that contain “parts” of words, to which we feel that we must assign meaning too. For example, the word *repaint* means something like ‘to paint again’. We therefore observe that the

word part *re-* adds some meaning of repetition to the meaning of the verb *paint*. Similarly, the final *-s* in the word *houses* adds the meaning ‘more than one’ to the noun *house*, i.e. it builds a plural noun. Forms like *re-* or *-s* are clearly meaningful, just like *cat* and *house* are meaningful but, by our definitions of “word” in section 5.2 above, they cannot be said to be words of English.

On the other hand, many words of English cannot be split into smaller units: *cat* and *house* are examples. The same is true of the word parts *re-* and *-s*. In order to account for the common properties of linguistic units like *cat* as well as *re-*, we need the concept of ‘a meaningful unit that contains no smaller meaningful units’. This concept is traditionally labelled a **morpheme**, duly defined as a minimal unit of meaning, i.e. the smallest meaningful unit in a language. By this definition, the words *cat* and *house* contain one morpheme each, as do the forms *re-* and *-s*, and words like *repaint* and *handbag* contain two morphemes each. By the same definition, two working assumptions follow:

- Any word consists of at least one morpheme, given that words express meaning, and morphemes are the smallest units of meaning in a language.
- Any morpheme must contribute meaning to the overall meaning of the word of which it is a part.

Morphemes are therefore **compositional** units of meaning. That is, words must be exhaustively broken up into morphemes. Let’s practise with a few examples.

#### Activity 5.11

How many morphemes would you say are contained in

words like the following?

caterpillar unhappier carpet cat-food uncle

We immediately observe that, as predicted, it is **meaning**, not **spelling**, that provides the guidelines for morpheme analysis. If spelling mattered, we would be justified in splitting the word *caterpillar* into *cater* and *pillar*, two well-formed words of English. The point is that neither *cater* nor *pillar* contribute to the meaning of *caterpillar* – nor does *cater-* in *caterpillar* sound like the word *cater*.

Similarly, the word *carpet*, which may sound like the words *car* and *pet* pronounced in sequence, does not obviously mean something like ‘a pet to be kept in a car’. That is, the meaning of *carpet* has nothing to do with cars or pets: *carpet* contains one morpheme only.

In contrast, the word *unhappier* clearly contains three morphemes, *un-*, *happy*, and *-er*. Note that, as usual, we are concerned with sound, not spelling: spelling *happy* with ‘y’ or ‘i’ is irrelevant for our analysis. All three morphemes contribute to the meaning of this word. The meaning of *happy* is part of the meaning of *unhappier*, *un-* contributes the negative meaning of the word, and *-er* contributes its comparative meaning, as discussed above for the word class Adjective.

On the pattern of *unhappier*, we might want to split *uncle* into *un-* and *cle*. This is clearly wrong, in that *cle* is meaningless in English and therefore not a morpheme. It follows that *un-* is not a morpheme either in this word, because *un-* cannot be contributing meaning to something that is itself meaningless. We then conclude that the word *uncle* contains one morpheme only.

Finally, the word *cat-food* contains two morphemes, that happen to be represented in writing with a hyphen in between (which is not

the case in e.g. *handbag*). The meanings ‘cat’ and ‘food’ are part of the meaning of the word *cat-food* (just like the meanings ‘hand’ and ‘bag’ are part of the meaning of the word *handbag*).

#### Activity 5.12

Can you explain this “Funny Dictionary” definition?

*Coffee – someone who is coughed upon.*

Which other words of English helped your reasoning?

Depending on the number of morphemes that they contain, words can be classified as:

- **Simple** words, that contain one morpheme only. For example, *clever, chimpanzee, a, from, the, mother.*
- **Complex** words, that contain more than one morpheme. For example, *establishment, impressive, kindness, karaoke-singers.*

Depending on their patterning, morphemes can in turn be classified as:

- **Free** forms, that are themselves words of the language. For example, *in, of, a, berry, intelligent, hand, bag.*
- **Bound** forms, that always occur as part of a word. For example, *re-, un-, -ly, -ness, -ish.*

### Activity 5.13

1. Explain whether the following words are simple or complex:

*rubbish*

*shoulder*

*girlish*

*friendship*

*party-goer*

*harness*

2. Classify the morphemes in these words as bound or free.

Section 6.3.2 in the next chapter provides a schematic summary of different types of words according to their morphology.

### *Morphemes and morphs*

Morphemes are abstract entities. They are constructs that we assume to exist in the system of a language, to help us describe features of that language.

Morphemes cannot therefore be spoken or heard. What can be spoken and heard is the concrete pronunciation of each morpheme, that we call a **morph**. To clarify this distinction between abstract concepts and their concrete realisations, let's discuss something that we all understand very well: food.

Supposing you've eaten chilli crab before, and we ask you: "Do you like chilli crab?", you may answer yes or no. Let's take a look at what went on in your head when you read this question. Whatever your

reaction to this dish, you thought of “chilli crab” as the generic name of a dish. You did not think of any individual crab cooked in chilli. Similarly, whether you love or hate classical opera, you don’t love or hate one particular opera but the generic type of music called “classical opera”. In other words, you have in your head generic concepts of chilli crab and classical opera.

Generic concepts are abstract entities: they are ideas that live in your mind. These ideas, however, are based on your concrete experience of, for example, chilli crab in real life: you’ve tasted it, and formed an opinion about it. The same is true of morphemes and morphs. Just like your different experiences of chilli crab helped you form the abstract concept of it, so our different experiences with hearing language spoken around us help us form abstract concepts about its units. We all pronounce and hear things differently, because we are all individuals. But we are all able to recognise the same units, no matter how differently pronounced they may be. For example, the song *Happy Birthday to You* is always the same song whether we, you, your neighbours or the National Chamber Choir sing it. Despite different renditions of it, we still recognise it as the song *Happy Birthday to You*. Similarly, the morpheme {banana} remains the morpheme {banana} whether you or someone else says it, whether we yell it or whisper it, and whether you say it in good health or with a badly blocked nose (in print, morphemes are usually represented between curly brackets, as we show here).

#### Activity 5.14

Ask three or four people to write the sentence I love chilli crab on a piece of paper. Then make it clear to yourself why they all wrote the “same” sentence, despite their different handwritings.

Since we are dealing with *spoken* language, let's see how these notions of concrete difference versus abstract "sameness" apply to the pronunciation of morphemes. Some morphemes are always pronounced in the same way, e.g. {un-}, {happy} and {-er}. However, the pronunciation of other morphemes sometimes varies, and we need to understand why this is so.

We said above that you could work out your own distribution by noting down all the places where you can be found. Supposing that you are a student with a keen interest in swimming and basketball, it is likely that we will find you at school, at the pool and in a basketball court. But this also means that you will look different in these three places, not least because you will be dressed in a way that is appropriate to each environment. Nevertheless, the way you dress is not central to your identity: you, in swimming gear, basketball kit or city clothes, are still the same person, not three different people. In other words, we would be able to recognise you as 'you' regardless of the way you actually look in these different contexts that are part of your distribution, because we know that your dress depends on the particular context where you happen to be at different times. The same is true of some linguistic units: they also dress appropriately to their contexts, as it were, and therefore look different despite being the same unit. Let's check with some data:

(5.9)

Set 1	Set 2
<i>an apple</i>	<i>a house</i>
<i>an owl</i>	<i>a tree</i>
<i>an idea</i>	<i>a mistake</i>

We observe that there is one word, *an*, preceding the nouns in Set 1, and a different word, *a*, preceding the nouns in Set 2. Before dismissing this puzzle as a random quirk of English, we may instead try to find a reason for it. Knowing now that context plays role in linguistic analysis, we reason that the cause of variation may lie in

some difference in the units that *follow* the forms *a* and *an*. The preceding context cannot provide any explanation, because it is the same: in both cases there is nothing, i.e. there is silence. We observe that the words following *an* in Set 1 all begin with a vowel (usually spelt with the letters *a, e, i, o, u*), and those following *a* in Set 2 all begin with a consonant.

We can then conclude two things. First, that the reason for the observed variation in form must be that the immediate context of each form is a vowel or a consonant, respectively; and second, that the two forms *an* and *a* must therefore represent two variants of one same unit, not two different units. They are two morphs of the same morpheme. We may now give the morpheme any label of our choice. Traditionally, as you may know, the label for this morpheme is {indefinite article}.

#### Activity 5.15

Our conclusion that the morpheme {indefinite article} has two different morphs was based on a very limited dataset and must therefore be provisional. But conclusions must be general, in order to have any scientific validity. In this case, our conclusion must predict other occurrences of *an* and *a*, given similar contexts. Our next step is therefore to hypothesise, from these observations, that *an* precedes a vowel and *a* precedes a consonant, and check this prediction with data that are not part of the original datasets.

Ask yourself which morphs of {indefinite article} occur in contexts such as these:



<i>cat</i>	<i>essay</i>	<i>understatement</i>	<i>salad</i>
<i>armchair</i>	<i>opportunity</i>	<i>van</i>	<i>necklace</i>

On the basis of these additional observations, is our hypothesis confirmed or disconfirmed?

### *Criteria to identify morphemes*

Given that certain morphemes always have the same morph, whereas others have different morphs, how do we know that we are dealing, at all times, with “the same” morpheme? Three criteria help us decide.

- **Sound.** A morpheme either has a constant pronunciation, or its pronunciation is predictable from its context. The morpheme {cat} is pronounced in the same way regardless of context, whereas the pronunciation of the morpheme {indefinite article} as *an* or *a* varies according to the context in which it occurs and can be predicted from it.
- **Grammar.** A morpheme regularly patterns with the same type of units. The morpheme *re-* in the word *repaint* consistently precedes verbs. The morpheme {indefinite article} can precede nouns or adjectives.
- **Meaning.** A morpheme has a constant meaning. The morpheme *re-* in *repaint* always means repetition of the action indicated by the verb, and the meaning of *an* vs. *a* remains the same, regardless of form.

Activity 5.16

Let's check these criteria against some data.

In the following set of words, can you identify one or several *-er* morphemes? Why?

singer	oyster	baker
greater	potter	louder

Try to work this puzzle out on your own before reading our analysis below! Explain your reasoning very clearly.

The first observation is that all forms spelt *er* are pronounced in the same way. In case all of them are found to be a morpheme, or the same morpheme, they therefore obey the sound criterion.

The best way to make sense of the two remaining criteria is to **paraphrase** the meaning of each word. A paraphrase uses different words to give the same meaning, describing it as clearly as possible. So what is a singer? The meaning of *singer* can be paraphrased as 'someone who sings'. By using the verb *sing* in the paraphrase we immediately realise three things that will help us solve our puzzle: one, the word *singer* contains the morpheme *sing*; two, this morpheme is a verb; and three, *-er* must be a morpheme too, because *singer* and *sing* are both meaningful words and they mean different things. We have thus found out that in the word *singer*,

-er contributes the meaning ‘someone who sings’, i.e. ‘someone who does the action represented by the verb’. The form -er is therefore attached to a verb. We can find a similar patterning in the word *baker*, where the verb *bake* plus the form -er also mean ‘someone who does the action represented by the verb’. The criteria of meaning and grammar are therefore satisfied: -er is the same morpheme in both words.

In the word *potter*, however, -er contributes a different meaning: *pot* can be a noun or a verb, but even taking it as a verb, the word *potter* does not mean ‘someone who pots’. It means ‘someone who makes pots’, where *pots* is a noun. The morpheme -er is therefore a different morpheme in this word, first because it attaches to a noun, and second because it gives the meaning ‘someone who makes the objects represented by the noun’ to the word *potter*.

The words *louder* and *greater*, in turn, show another pattern: -er attaches to an adjective, contributing the comparative meaning ‘more of the quality expressed by the adjective’. The word *oyster* clearly contains one morpheme only: *oyst* is meaningless, and so is *er*, in this word.

We then conclude that the data show examples of three different morphemes, that all happen to be pronounced in the same way. There is one -er morpheme attaching to verbs, with the shorthand meaning ‘someone who Verbs’ (or shorter still, ‘someone who Vs’), another -er morpheme attaching to nouns, meaning ‘someone who makes N’, and a third -er morpheme attaching to adjectives, meaning ‘more Adj’.

One reminder: our linguistic analysis concerns spoken language, i.e. **speech**. Discrepancies in **spelling** like *sing-singer* vs. *bake-bak(e)r* should be disregarded throughout this book. Both *sing* and *bake* end in a consonant sound, despite their spellings.

#### Activity 5.17

Can you explain this lame joke?

Question. *Where does a general put his armies?*

Answer. *In his sleeveies.*

Which other words of English helped your reasoning?

### *“Meaningless” morphemes?*

The meaning compositionality of certain words is a matter of controversy. How many morphemes would you count in words like *strawberry* or *ladybird*?

The morphemes *straw*, *berry*, *lady* and *bird* exist in English, but they do not contribute any meaning to these two words that they apparently form. True, a strawberry is a berry, but how do we fit the meaning of *straw* into the overall meaning of the word *strawberry*? By the approach taken in this chapter, morphemes are compositional units of meaning. This entails that if we cannot assign *compositional* meaning to word parts, even if those word parts are words of the language in their own right, with clear meanings, then we must conclude that those word parts are not morphemes.

A similar problem arises with words like *gooseberry* and *cranberry*. The first part of these *berry* words, although non-compositional (or simply meaningless, like *\*cran*), nevertheless appears to serve the function of distinguishing one type of berry from another. This may be so, but the assumptions that we made for morpheme analysis concern the meaning of morphemes, not their distinctive function

in words. We are thus forced to analyse *berry* words like these as simple words, and to do the same for words like *ladybird* or *butterfly*. The word *cranberry*, incidentally, gained fame in morphological analysis because it became the technical term for words like itself: the apparently meaningless word parts of words that appear to contain one other genuine morpheme are called **cranberry morphemes**.

As in our earlier discussion of adverbs, here too we see that particular analytical frameworks, which necessarily include assumptions, sometimes leave our analyses with several loose ends. Alternative frameworks, with different assumptions, will reach different conclusions. This is as true, and as natural, in the science of language as in other sciences: physicists, for example, are also divided about whether it makes more sense to talk about light as consisting of waves or of particles, to mention just one of many controversies in physics. This is where the controversy lies: in any science, controversy simply means different ways of looking at the same things. Following the assumptions introduced in this chapter, we next look at how morphemes combine in different ways to form different types of words.

## Food for thought

*Why English is so hard to learn*

We must polish the Polish furniture.

He  
could  
lead if he  
would  
get the  
lead out.

The  
farmer  
used to  
produce  
produce.

The dump was so  
full that it had to  
refuse more refuse.  
The soldier decided  
to desert in the  
desert.

This was a good time to present the present.

A bass was  
painted on the  
head of the  
bass drum.  
When shot at,  
the dove dove  
into the  
bushes.

I did not object to the object.

The  
insuranc  
e was  
invalid  
for the  
invalid.

The  
bandage  
was  
wound  
around  
the  
wound.

There was a row

among the oarsmen  
about how to row.  
They were too close  
to the door to close it.

The buck does funny  
things when the does  
are present. They sent a  
sewer down to stitch  
the tear in the sewer  
line. To help with  
planting, the farmer  
taught his sow to sow.  
The wind was too  
strong to wind the sail.

After a  
number of  
injections my  
jaw got  
number. Upon  
seeing the tear  
in my clothes I  
shed a tear. I  
had to subject  
the subject to a  
series of tests.

How can I intimate this to my most intimate friend?

## Further reading

Deterding, David H. and Poedjosoedarmo, Gloria R. (2001). Chapter 3. Word classes. In *The grammar of English. Morphology and syntax for English teachers in Southeast Asia*. Singapore: Prentice Hall, pp. 18-35.

Hudson, Grover (2000). Chapter 4. Morphemes. In *Essential introductory linguistics*. Oxford: Blackwell, pp. 57-68.

## Reference

Sapir, Edward (1921). *Language. An introduction to the study of speech*. New York: Harcourt Brace.

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.



# Chapter 6: Morphology- Word formation

## The grammar of words: word building

### Chapter Preview

How are new words formed?

What kinds of new words can be formed?

What does word formation tell us about the grammar of words?

What does the grammar of words tell us about the grammar of language?

## 6.1 Introduction

Speakers keep their languages alive and usable by changing the vocabulary of their languages (and, less easily, their grammar) according to what they need to express.

Language users do this in three major ways. One way is to simply import a useful word from another language, just like people import useful products from other countries. This is how an Italian word like *pizza* or a Japanese word like *karaoke* became English words. Words that are circulated in this way among languages are called **borrowings**. Secondly, language users can change the meaning of words already in the language, to make them mean different things. The English word *sad*, for example, is currently used to mean something similar to 'pathetic', besides keeping its meaning of 'unhappy'. In this new use, a *sad joke* is not a joke that makes you cry, but a joke that doesn't make you laugh. A third way of creating new words in a language involves manipulating not just their meaning but also their grammar, by disassembling the morphemes from the words in which they appear, and reassembling them into new words. This is what word formation is about. Knowing that a morpheme *-er* means 'someone who Vs', as we saw in the previous chapter, we can safely create a brand-new noun *emailer* to mean 'someone who emails', and use it straight away to say that *My friend Janice is a compulsive emailer*. Speakers of English will have no difficulty understanding what this new word means, even if they have never heard it before, given their knowledge of the meaning and the grammar of the morphemes making up this newly-minted word. Whether the word will ever become accepted in English is another matter.

This chapter deals with several processes that allow language users to build new words, often called *word formation processes* in the literature, with special emphasis on processes that involve the grammatical make-up of words.

## 6.2 Word formation

**Word formation** concerns the processes that allow us to create new words with grammatical resources already available within a language. These processes must of course obey the rules of the language, i.e. its grammar. The word *emailer* is a well-formed word of English, as are other possible words like *downloader* or *rebooter*, because they follow the same word-formation rule of English that allows words like *writer* or *daydreamer*.

### Activity 6.1

Can you explain why this ad for an entertainment outlet is an example of creative language use?

*Never Unfun*

Assuming, as we did in the preceding chapter, that words are made up of morphemes, word formation involves a patterning of morphemes within words, whose rules we can find out. Let's use some data to see what we mean by morpheme patterning.

(6.1)

---

<i>room</i>	<i>rooms</i>	<i>*sroom</i>	<i>*roosm</i>
<i>darkroom</i>	<i>*roomdark</i>	<i>darkrooms</i>	<i>*darksroom</i>
<i>unhappy</i>	<i>*happyun</i>	<i>happily</i>	<i>*lyhappy</i>
<i>commit</i>	<i>commitment</i>	<i>commitments</i>	<i>*commitment</i>
<i>songbird</i>	<i>birdsong</i>	<i>*unments</i>	<i>*mently</i>

---

We observe that:

- Morphemes must occur in certain positions within a word. For example, the {plural} morpheme in *rooms*, spelt -s, must occur at the end of the word, not at the beginning (*\*sroom*) or in the middle of it (*\*roosm*).
- The word class to which lexical morphemes belong is important for their ordering within complex words. The forms *darkroom*, *songbird* and *birdsong* are well-formed, whereas the form *\*roomdark* is not.
- Certain bound forms must occur before others. The form *commitments* is acceptable, whereas the form *\*commitment* is not.
- Bound forms cannot be combined with one another. Whereas *birdsong*, with two free forms, is well-formed, neither *\*unments* nor *\*mently* are.

Observations like these help us tell apart different types of morphemes, which in turn helps us tell apart different word formation processes.

There are two players involved in word formation processes. If we think of building words as we think of, say, building a wall, we need

the items that we are going to put together (morphemes, or bricks of various types) and we need a way of putting them together (rules, or a building plan). Attempting to fit morphemes together at random won't result in words, just like throwing bricks around or heaping them together won't build a wall. We need both building blocks and constraints to build walls and words properly.

In word formation, the building blocks are of two types, and so are the constraints.

### *The building blocks*

- **Stem:** a morpheme, or a word, to which other morphemes can attach.
- **Affix:** a morpheme that attaches only to a stem.

How can these two concepts help us explain some of the observations above? We can see, from the data in (4.1), that the words *commitment* and *happily* are well-formed, whereas *\*mently* is not. Using the concepts just introduced, we can now explain why this is the case. Both *commitment* and *happily* are complex words, i.e. words comprising more than one morpheme. Both words also comprise a stem and an affix: *commitment* comprises the stem *commit* to which the affix *-ment* attaches, while *happily* comprises the stem *happy* to which the affix *-ly* attaches. In contrast, if we treat *\*mently* as a complex word, it seems to comprise two affixes (*-ment* and *-ly*) attached to one another, rather than a stem and an affix. But our definition of affix says that affixes only attach to stems, not to other affixes.

#### Activity 6.2

Go back to our data in (6.1), and do two things.

First, decide whether each of the building blocks in rooms, unhappy, darkrooms, songbird and \*unments is a stem or an affix.

Then, with the help of this decision, explain why rooms, unhappy, darkrooms, and songbird are well-formed, but \*unments is not.

### Activity 6.3

Knowing that un- and -able are both affixes in words like uncomfortable, explain whether the underlined word in the following sentence is a counterexample to the rule that affixes can attach only to stems:

*Small children are quite unable to keep still for long periods of time.*

There is another technical term used to refer to the fundamental stem, as it were, of a word. In the word *commitments*, this stem is *commit*, the basic word from which the complex word *commitments* is built. We then say that the root of the word *commitments* is *commit*. The **root** of a complex word is itself a word from which all affixes have been removed. We can visualise this word formation process as follows, where the arrow indicates the result of word building:

(6.2)

*commitments*

root *commit* + affix -ment → *commitment*

stem *commitment* + affix -s → *commitments*

This example shows that a root can be a stem, but that not all stems are roots.

### Activity 6.4

How many stems does the word *disgraceful* contain?

Write them down and explain how you reasoned to find them.

### *The constraints*

- **Hierarchy:** the internal structure of complex words is hierarchical.
- **Well-formedness:** each step in word formation must produce a well- formed word of the language.

These two constraints help us make sense of word formation. Going back to our analogy of building a wall, they reflect the commonsense observation that walls are built layer by layer, and that each brick added to a wall in fact builds a small wall of its own by fitting neatly among its neighbours. We follow a similar reasoning with word building: complex words are built up step by step from stems and/or affixes, and each intermediate word must itself be a well- formed word. As shown in example (6.2), a word like *commitments* is formed by attaching the affix *-ment* to the root/ stem *commit*, forming the word *commitment*, a new well-formed stem to which *-s* in turn attaches. In addition, knowing that *dark* is an Adj and *room* is a noun in the complex word *darkroom*, and that Adj precedes N in English, we can explain why *darkroom* is well-formed whereas *\*roomdark* is not.

## 6.3 Major word formation processes

The most productive word formation processes in English are **affixation**, **compounding** and **conversion**, the ones that we deal with in greater detail in this chapter.

In morphology, **productivity** means the degree to which a word-formation process is used in a language. We might use an analogy of productive worker bees – the most productive worker bee is the one that makes the most honey. So also the most productive word formation rules are the ones that are used most frequently to create new words in a language or language variety. Generally, productivity is directly proportional to **compositionality**, the degree to which the meaning of a new word is predictable from the meanings of its constitutive morphemes. That is, “more productive” entails “more compositional”, and vice versa. For example, an affix like {plural} –s is extremely productive, in that new nouns in English can be made plural by using it. It is also compositional, in that it consistently contributes the meaning ‘more than one’ to the new word. If *emailer* is a noun, then *emailers* is its plural.

However, we should note that compositionality is not an absolute matter. It is not the case that processes, or words meanings, are either compositional or non-compositional. Rather, compositionality is understood as a cline: at one end of this cline, we find **transparent** word meanings which are easily deduced from the meanings of the morphemes that make up the word; at the other end of this cline, there are **opaque** meanings which are not easily inferred from the morphemes making up the word. We will see below several examples of degree in compositionality.

### 6.3.1 Affixation

Affixation is one of the most productive word formation processes



in English. In **affixation**, an affix attaches to a stem. All the words in the sentence *Teachers dislike yawning students* are affixed words. We can analyse affixes based on two criteria: according to their distribution, and according to their meaning.

### *Distribution of affixes*

We said above that affixes must attach to a stem, but we did not clarify the order of attachment of stem and affix. We now add that there are different types of affix, according to their distribution. For example:

- **Prefixes** precede the stem.
- **Suffixes** follow the stem.

We need this clarification in order to explain why the two words *unhappy* and *happily* in (6.1) are well-formed, whereas *\*happyun* and *\*lyhappy* are not: *un-* is a prefix, and *-ly* is a suffix. That is, *un-* must precede the stem to which it attaches, while *-ly* must follow its stem. Note the use of a dash following or preceding these affixes. This is essential to make clear whether we're referring to a prefix or a suffix.

These two types of affix account for affixation processes in English. Other languages have additional types of affix. For example, both Bontoc, a language spoken in the Philippines, and *Tukang Besi*, an Austronesian language spoken in Indonesia, have *infixes*, affixes that appear in the middle of a stem. Malay, meanwhile, has *circumfixes*, affixes that surround a stem, in addition to prefixes and suffixes, as illustrated in Figure 6.1.

<b>Bontoc</b>		
Adj/N		Infix
<i>fikas</i> 'strong'	<i>fumikas</i> 'to be strong'	-um-
<i>fusul</i> 'enemy'	<i>fumusul</i> 'to be an enemy'	-um-
<b>Tukang Besi</b>		Infix
Adj/V		
<i>to'oge</i> 'big'	<i>tumo'oge</i> 'biggest'	-um-
<i>tinti</i> 'run'	<i>tuminti</i> 'running'	-um-
<b>Malay</b>		
Adj	N	Circumfix
<i>selamat</i> 'safe'	<i>keselamatan</i> 'safety'	<i>ke-___-an</i>

Figure 6.1. Examples of affixes in Bontoc, Tukang Besi and Malay

### Activity 6.5

Do any of your other languages have infixes or circumfixes?

### Meaning of affixes

According to meaning, affixes can be of two types.

- **Derivational** affixes form a new word with a new lexical meaning.
- **Inflectional** affixes form a variant of the word they attach to, adding a grammatical meaning.

You will notice that this difference in the kinds of meanings conveyed by affixes parallels the difference that we discussed in the previous chapter, concerning lexical and grammatical words. Like lexical words which express ideas/concepts, derivational affixes have semantic content. Derivational affixes are so named because when they attach to a root/stem, they derive a new word, i.e. a word with a new lexical meaning. In contrast, inflectional affixes, like grammatical words, carry grammatical meaning. They mark grammatical properties such as tense, number, person and case, and do not change the lexical meaning of the words they attach to.

This difference between lexical and grammatical meaning explains why certain words are regularly given an entry of their own in dictionaries, whereas other words share the same entry. For example, the words *commit* and *commitment*, though related, are in fact two words, with two different lexical meanings that entitle each to a separate dictionary entry. In contrast, inflected words (e.g. *rooms*) are listed under the same entry as their root, given that they represent grammatical variants of the same word.

In derivational affixation (or **derivation**, for short), the word class of the stem and the word class of the derived word may or may not be the same. This means that derivational affixes may be **class-maintaining** or **class-changing**. Consider these two words:

(6.3) *unhappy*      *commitment*

Affixing *un-* to the Adj *happy* derives a new Adj (*unhappy*); *un-* is a class-maintaining derivational affix. Affixing *-ment* to the verb *commit* also derives a new word (*commitment*), but this time the lexical class of the derived word changes to a noun; *-ment* is a class-changing derivational affix.

Inflectional affixes, as we saw in (6.1), change the grammatical meaning of the words they attach to. Consequently, inflectional affixation (or **inflection**) is always class-maintaining. For example, inflectional affixation with plural *-s* changes the grammatical meaning of the singular noun *room* to plural *rooms*, but the lexical category remains unchanged. Both *room* and *rooms* are nouns. Similarly, affixation with *-ed* changes the grammatical meaning of

walk from present tense to past tense *walked*, but the lexical category remains unchanged. Both *walk* and *walked* are verbs. If we assume that lexical meaning is more central than grammatical meaning, we can see why inflectional affixes regularly follow derivational affixes in the formation of words. One example is the word *commitments*, discussed in section 6.2 above.

Using the two criteria of distribution and meaning, we can distinguish English affixes in the following way:

	derivational	inflectional
prefix	√	X
suffix	√	√

Figure 6.2. Types of affix in English

Figure 6.2 shows that the derivational affixes of English can be either prefixes or suffixes. For example, *un-* in *unhappily* is a derivational prefix, while *-ly* in the same word is a derivational suffix. In contrast, the inflectional affixes of English are all suffixes. In fact, contemporary English has only eight inflectional affixes: four bound to verbs, two bound to nouns, and two bound to adjectives.

Activity 6.6

Can you identify the eight inflectional affixes of contemporary English?

Other languages, however, have inflectional prefixes as well as suffixes. One example is Swahili (the major African lingua franca). In many languages, nouns must belong to different grammatical

classes, called *genders*. You may be familiar with gender from languages like French, which has two (masculine and feminine) or German, which has three (masculine, feminine and neuter). In these languages, gender is marked by suffixes. Swahili has several different genders (e.g. for ‘human’, ‘other living things’, ‘liquids’, etc.) and all are marked with inflectional prefixes. As in other gendered languages, adjectives qualifying a noun must show the same gender inflection as the noun. Here is one example from Swahili with a noun and an adjective for the gender sometimes called “Class 6” (in Swahili, the adjective follows the noun):

Figure 6.3. Example of gender inflection in Swahili

Swahili

N	Adj	Inflectional prefix
tunda ‘fruit’	zuri ‘good’	ma-
matunda mazuri	‘good fruit’	

*Affix identification*

In section 5.4.3, we listed three criteria to identify morphemes, namely sound, grammar and meaning. The same criteria can of course be used to identify different affixes. Let’s see how the three criteria apply to the affix *-ly* in the words *sharply*, *kindly* and *happily*. If all three criteria are obeyed, then we are dealing with the

same affix. If any one criterion is not met, then we are dealing with different affixes.

- **Sound.** In all three words, the affix is pronounced the same way, [li] (conventions to represent pronunciation are dealt with in the next chapter). In other cases, the pronunciation of an affix may be predictable by rule. Predictable variation of this kind also satisfies the sound criterion.
- **Grammar.** The affix attaches to the same stem class, in the same position (as a prefix or as a suffix), and the lexical category of the resulting word must be the same for all the words under consideration. In this case, *-ly* is suffixed to an Adj to form an Adv in all three words.
- **Meaning.** The affix establishes a regular meaning relationship between the stem and the word resulting from the affixation. In this case, the meaning of the newly derived word can be paraphrased as ‘in a manner’, where the Adj replaces the blank. For example, *sharply* means ‘in a *sharp* manner’.

Note that the meaning **paraphrase** must contain the stem of the word, in this case the adjectives *sharp*, *kind* and *happy*, in order to make the meaning relationship between the stem and the derived word absolutely clear. We can now generalise our observations about the formation of the words *sharply*, *kindly* and *happily* to all other words containing the same affix by means of a shorthand rule, like this:

(6.4) Adj + *-ly* → Adv, ‘in a \_ manner’

In **rule notation** of this kind, the plus sign represents sequential ordering of morphemes, and the arrow indicates the result of that ordering. This rule summarises all the information that we need, in order to identify the affix *-ly*. You can use this rule to check for yourself that *sharply* contains the same affix as words like *brightly*, *lightly* or *beautifully*.

The observations and analysis that we developed in this section

of course apply to any complex word formed through affixation, not just the three adverbs under discussion here. Otherwise, our conclusions would be useless in a scientific account of language.

#### Activity 6.7

Use rule notation to account for the formation of the words *unhappy* and *commitment*. Then find three other words that follow the same rules.

### 6.3.2 Compounding

Affixation involves attaching one or more affixes to a stem. In contrast, **compounding** involves attaching a stem to another stem. In the following sentence, the words in italics are compounded words:

(6.5) Janice spilled the *salad dressing* on her *brand-new laptop*.

Notice that spelling is irrelevant for the identification of compounds. Compounds may be spelt with hyphens as in *brand-new*, without hyphens as in *salad dressing*, or as single words as in *laptop*. What is crucial is the meaning relationship between the stems making up the compound word.

#### *Form of compounds*

Each of the stems in a compound is itself a word of the language, and therefore belongs to a particular word class. However, the word class of each stem does not necessarily correspond to the word class of the compound word itself, as shown in the table below:

Stems	Compound	Word class of stems	Word class of compound
<i>hand + bag</i>	<i>handbag</i>	N + N	N
<i>pick + pocket</i>	<i>pickpocket</i>	V + N	N
<i>pull + over</i>	<i>pullover</i>	V + P	N
<i>sea + sick</i>	<i>seasick</i>	N + Adj	Adj
<i>bare + foot</i>	<i>barefoot</i>	Adj + N	Adj
<i>run + down</i>	<i>rundown</i>	V + Adv	Adj
<i>spoon + feed</i>	<i>spoonfeed</i>	N + V	V
<i>over + shadow</i>	<i>overshadow</i>	P + V	V
<i>in + to</i>	<i>into</i>	P + P	P

Figure 6.4. Word classes of compounds and their stems

Figure 6.4 shows that the word class of the compound word may be the same as the word class of one of its stems, often the right-hand stem (as in *pickpocket*, *seasick*, *spoonfeed*, and *overshadow*), but that this need not always be the case (as in *barefoot*, *rundown* and *pullover*). There is wide variability in the correspondence of word class between stems and compound, and it is this flexibility that contributes to the lively productivity of compounding as a word-formation process.

Languages like English allow simple juxtaposition of stems to form a compound, as in the examples above. This is the commonest compounding process in these languages. But other compounding processes exist, such as linking stems by means of grammatical words as in *mother of pearl*, *chief of staff* or *black and white*. Examples are expressions like *a mother of pearl necklace* or *a black and white photograph*. Other languages prefer linking stems in this way, for example Romance languages like French or Portuguese.



What's important is that the words so linked, whether by simple juxtaposition or through the use of linking words, acquire a specific meaning of their own, that is different from the meaning of each of the stems that make up the compound. Compound words, like derived words, have dedicated entries in dictionaries.

#### Activity 6.8

We saw in Activity 6.4 that the word *disgraceful* contains more than one stem. This being so, explain why this word is not a compound.

#### *Meaning of compounds*

A compound word encapsulates a specific concept. The meaning of many compounds is non-compositional and may lie anywhere from transparent to opaque on a compositionality cline. What this means is that we cannot predict the exact meaning of a compound by assuming a particular relationship between the stems that build it. Take, for example, the compounds *meatball* and *handball*. They both have the structure  $N + N \rightarrow N$ , and they both mean something that is related to *meat* and *ball* in the first case, and to *hand* and *ball* in the second. But whereas *meatball* means 'a ball made of meat', *handball* doesn't mean 'a ball made of hand(s)'. Similarly, *handbag* means 'a bag to be carried in your hand', whereas *handball* does not mean 'a ball to be carried in your hand'.

### Activity 6.9

Explain why this is an example of language play:

*If olive oil is made by pressing olives, how is baby oil made?*

As illustrated above, the meaning of some compounds is opaque because of the idiosyncratic relationships between the stems forming the compound. But compound opacity can also result from meaning shifts in the stems of a compound. The compound *blackboard*, for example, was created at a time when all school boards were black, i.e. the stem *black* was used in its literal sense. Nowadays, however, we can talk about *green blackboards* and even about *white blackboards* (although the word *whiteboard* has been coined for the latter), without feeling that we are being paradoxical about the colour of the board. The reason is that the stem *black* no longer designates the colour 'black' in this compound. Together with the stem *board*, it identifies a particular kind of object instead.

Despite the opaque meaning that the first stem contributes to these compounds, there is a sense in which compounds like *blackboard* or *darkroom* do have a transparent meaning, in that a blackboard is a board, and a darkroom is a room. The same cannot be said of the meaning of compounds like *pickpocket* or *pullover*. In compounds of the former type, the second stem is central to the meaning of the whole compound. We can paraphrase the meaning of compounds like *handbag* or *shoulder-bag* by saying that they are bags of a particular kind. Similarly, we can paraphrase the meanings of compounds like *seasick* and *car-sick* by saying that they both involve being sick in some way. Compounds of this type are called **headed compounds**: the second stem is the **head** of the compound,

and the first is its **modifier**. Two properties can be observed among these compounds, relating to:

- **Meaning:** the modifier narrows down the meaning of the head.
- **Word class:** the compound word belongs to the same word class as its head.

By these two properties, the meaning of a headed compound can be said to refer to a *kind-of* the meaning of its head. For example, a handbag is a kind of bag (for more on *kind-of* relations between word meanings, see section 9.5.2). The meaning of these compounds tends to lie on the transparent segment of the compositionality cline, compared to the meaning of **non-headed compounds** like *pickpocket* or *pullover*.

The productivity of compounding is borne out by the frequency with which so-called **long compounds** are formed. Long compounds are expressions formed by successive compounding of other compounds (this kind of compounding is an example of **recursion**. In English, 3-word and 4-word compounds are very common. Two examples are, with their stems numbered for ease of reference:

(6.6)

a.	vehicle	breakdown	service	
	1	2	3	
b.	professional	children's	entertainment	troupe
	1	2	3	4

Because long compounds are formed by compounding other compounds, we need to take account of **hierarchy** in their formation. Often, decisions about the order in which the stems attach to one another result in quite different interpretations of the meaning of the final compound. We would all agree that example (6.6a) can only mean 'a service dealing with vehicle breakdown', not 'a breakdown service for vehicles'. That is, (6.6a) is formed by attaching stems 1 and 2 to each other, and then stem 3. Similarly,

(6.6b) can only mean ‘a professional troupe for children’s entertainment’, not ‘an entertainment troupe made up of professional children’). That is, (6.6b) is formed by attaching stems 3 and 4 to each other, followed by stems 2 and 1, in that order.

But how would we parse a long compound like *Singapore noodles soup*? Do we mean ‘a kind of soup with noodles that is served in Singapore’, or ‘a soup containing Singapore noodles’? The two interpretations can be made clear using square brackets for the stems that are parsed together, like this:

(6.7)

a. [*Singapore*] [*noodles soup*]

b. [*Singapore noodles*] [*soup*]

The choice of interpretation may well depend on what we understand a “concept” to be, in the sense discussed above for the meaning of compounds. *Singapore noodles* may be a concept for certain speakers, in which case the parsing in (6.7a) is the one that immediately comes to mind. For other speakers, the compound may be **ambiguous**, i.e. mean two different things according to the alternative analyses in (6.7).

An illuminating episode concerning compound parsing involved one of the authors of this book as main character. As a newcomer to Asia, I saw a poster describing a *red dragon boat team*. My first reaction was to wonder “What colour is the dragon??” Would you have any trouble assigning a colour to the dragon too, or would you find the issue irrelevant? The explanation for my confusion is that, for me, *dragon boat* was not a compound concept, so I didn’t know whether to parse *red dragon* first, or *dragon boat* first. Think for yourself how you would parse a long compound like *kitchen towel rack*, which is always ambiguous because there is no single “basic” compound concept involved in its formation.

#### Activity 6.10

Do you find these long compounds ambiguous? Explain why you think so.

*busy family schedule*

*toy car factory*

*wooden door latch*

Here is a schematic summary of the word types discussed so far:

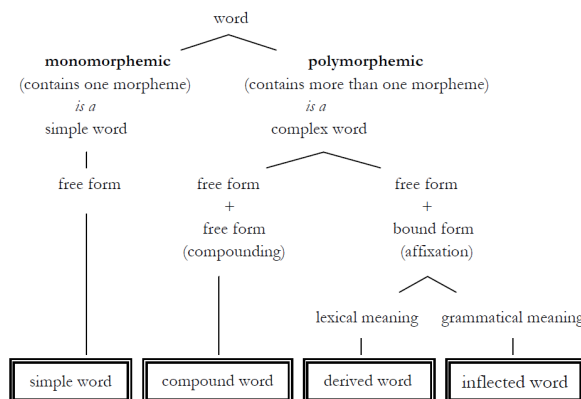


Figure 6.5.  
Summary of  
simple and  
complex  
words

### 6.3.3 Conversion

The last of the three highly productive word formation process that we wish to discuss here is conversion. **Conversion** involves a change in the word class of a word without any change in the form of the word. Examples of converted words appear in *italics* below:

(6.8) If you *bookmark* your favourite websites, they'll *cookie* every *download*.

Used originally as nouns, the words *bookmark* and *cookie* are

currently used also as verbs. The converse is true of the word *download*, which started life as a verb and is now used also as a noun. The productivity of conversion is seen in the vast number of identical word forms that serve as different word classes. A few examples include *judge*, *fast*, *party*, *impact*, and *email*. Out of context, the word class of converted words cannot be determined. In English, virtually any word can be converted to a noun. This is why we can talk about *the rich*, *a have-not*, or *the whys, ifs and buts* of an argument.

Conversion differs from both affixation and compounding in that new words are formed not through the addition of morphemes (whether affixes or stems) but simply by changing their word class. Because of this, conversion adds new simple words to the language, from other simple words, whereas affixation and compounding add complex words.

When dealing with word formation processes, it may sometimes be useful to find out the original word class of a word. **Etymology** (from the Greek *etymon*, ‘true meaning’ and *logos*, ‘science’) deals with the historical evolution of word meanings. Any good dictionary, such as the Oxford English Dictionary (OED) or the Merriam-Webster, for British and American uses of English, respectively, will provide this information. Incidentally, a lot of people think that the original meaning of words is their ‘true meaning’, as the Greek word above suggests. If so, we would all be very wrong in our current use of many words. Take *nice*, for example. At the time English imported this word from Latin through French in the Middle Ages, it mean ‘ignorant, foolish’ – certainly not how we use this word today.

## 6.4 Other word formation processes

We now discuss briefly four other word formation processes that are common in various languages. All of them share one

characteristic that sets them apart from the three major processes discussed so far: they all shorten words.

### 6.4.1 Backformation

**Backformation** is so named because it is the opposite of affixation: it involves removing from a word a part of it that is perceived as an affix. The word is taken “back”, as it were, to its stem “form”.

The interesting feature of backformed words is that the supposed affix is in fact not an affix at all, and there is therefore no stem to go back to. By analogy with other legitimately affixed words of the language, backformation in fact creates a new word. One classic example of backformation will help explain how it works. The word *television* was created as-is to designate what we all know it to mean. By analogy with pairs of words like *supervision-supervise*, *revision-revise*, the word *television* was (wrongly) assumed to be a derived word too, and the new verb *televise* was backformed from it. Many backformed words create verbs from nouns in similar ways. Examples include *hawk* from *hawker*, *edit* from *editor*, and *electrocute* from *electrocution*.

Backformation isn't always clear-cut, and at times may cause hesitation in the use of certain word forms. For example, when you find your bearings do you *orientate* or *orient* yourself? And are you then *orientated* or *oriented*?

### 6.2.2 Clipping

In contrast to backformation, **clipping** simply cuts a word short, without reference to morphological structure. Examples of clipped words include *exam* from *examination*, *maths* from *mathematics*, and *pub* from *public house*. Many students taking *English Language* refer to their course as *Elang*. Many of us surf the *net* rather than

the *internet*, and ride in *cars* rather than *motorcars*. These examples show that clipping can affect any part of the original word, its beginning, end or middle. The words *fridge* and *flu*, from *refrigerator* and *influenza*, for example, retain the middle, while clipping off the beginning and end of the words.

#### Activity 6.11

Can you explain the language play in the sign below, painted on the side of an electrician's van?

*Let us remove your shorts*

### 6.4.3 Acronymy

**Acronymy** involves using the initial letters of a sequence of words or morphemes to form a new word. We mentioned the word *television* above, as the name of a familiar object, but the likelihood is that you don't watch *television*, you watch its acronym TV instead. Other examples of acronyms are KL for *Kuala Lumpur*, MMR for (vaccination against) *measles, mumps, rubella*, DOS for *disk operating system*, or UNESCO for *United Nations Educational, Scientific and Cultural Organisation*.

These examples in fact conflate two types of acronyms. Some, like KL, are pronounced by the names of the letters that compose them, whereas others, like UNESCO, can be pronounced as a word. The former are sometimes called **initialisms**, whereas the latter are acronyms proper. The word CD-ROM is a mixture of both, its first part an initialism and its second part an acronym. Words like PhD



(*Philosophy Doctor*) or *radar* (*radio detecting and ranging*) are also taken as acronyms, although they both take two initial letters from one of their words, rather than just one (“Ph” from *Philosophy* and “ra” from *radio*).

Activity 6.12

In Singapore, the names of most expressways are shortened in the following way (all shortenings are read as initialisms).

Bukit Timah Expressway	BKE
Kranji Expressway	KJE
Pan-Island Expressway	PIE
Seletar Expressway	SLE

1. Can you find the rule for these shortenings?

2. Now try to predict the shortenings for the following expressways:

Central Expressway

Tampines Expressway

3. Think about naming practices of this kind in your own country, for roads, institutions, services, etc. Any interesting examples?

Once acronyms become words in their own right, they behave like ordinary words, exhibiting the features of the word class to which they are assigned. We can thus pluralise nouns like *radar* and CD-ROM, to talk about *radars* and CD-ROMs, respectively. Spelling, particularly of proper acronyms, also normalises to lowercase letters. This is the case for *radar*, as it is for *scuba* and *laser*, from *self-contained underwater breathing apparatus* and *light amplification by stimulated emission of radiation*, respectively.

These examples make it easy to understand why acronymy is an

economical way of using words, and why, therefore, acronyms are extremely common in any media where speed of communication is seen as desirable, e.g. chatrooms, email, instant messaging systems. Other recent examples of acronyms include SARS for *severe acute respiratory syndrome*, DVD for *digital video disc*, URL for *uniform resource locators* and SMS for *short message service*.

#### Activity 6.13

1. Collect a file of commonly used acronyms that you use on email or when messaging your friends.
2. Make a list of acronyms used in the news (either on TV/radio or in print).

### 6.4.4 Blending

A **blend** can be seen as the compounding of clipped words, in that it takes segments from words and joins them together in a new word that retains meaning characteristics from the original words. The word *smog*, for example, is a blend of *smoke* and *fog*, and means a ‘blend’ of smoke and fog. Similarly, *brunch* is a blend of breakfast and lunch, a *modem* is a blend of a modulator and demodulator, while a *dramedy* blends drama and comedy.

Other examples of blends are the names by which local varieties of languages are known. Examples of labels involving English include *Hinglish* (Hindi English), *Japlish* (Japanese English), *Swenglish* (Swedish English), and *Spanglish* (Spanish English). These blends reflect the dual contribution of their two referents to form the language variety in question. Within this set of labels, the word *Singlish* is also a blend, although its first clipping refers to a country (Singapore) rather than a language. The same applies to *Manglish* (Malaysian English), the variety of English spoken in Malaysia.

# 6.5 Morphological analysis of complex words

Several of the word formation processes discussed in this chapter can, and in fact do, operate on the same word. This flexibility is part of the productivity of these word formation processes. We can, for example, find words like *ATMs*, formed through acronymy and affixation, or like *piano-players*, where compounding and affixation apply. Let's now see how complex words like these are analysed.

## 6.5.1 Interpretation of meanings

Insight into the meaning of a complex word is best gained by means of a **paraphrase** that explicitly describes its meaning. Paraphrase makes clear not only the grammatical identity of each of the component morphemes in a complex word, but also the grammatical relations among them. As highlighted above, paraphrases must mention the morphemes that constitute a word, so that the meaning of the word becomes clear. For example:

Word	Paraphrase
<i>handbag</i>	a bag to carry in your hand
<i>barefoot</i>	with feet that are bare
<i>deep-fry</i>	to fry (something) deeply
<i>insane</i>	not sane
<i>piano-player</i>	someone who plays the piano
<i>piano-players</i>	more than one piano-player

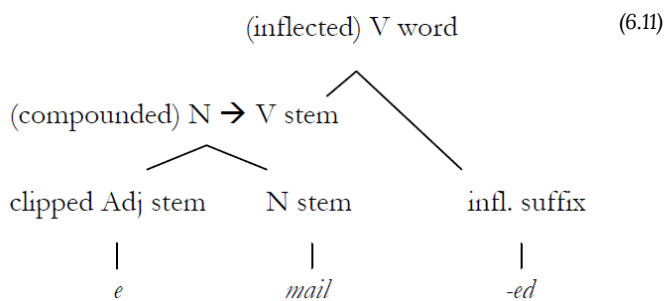
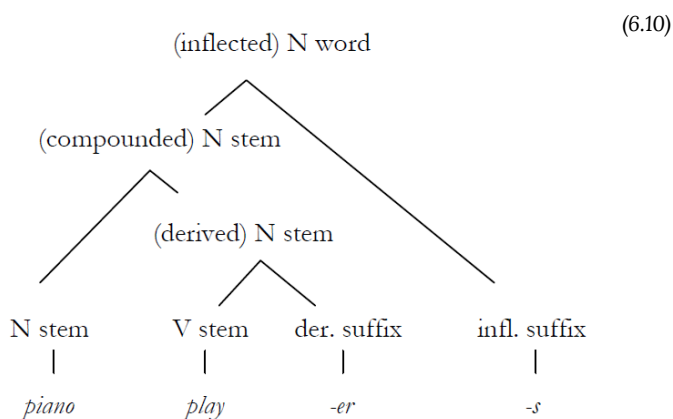
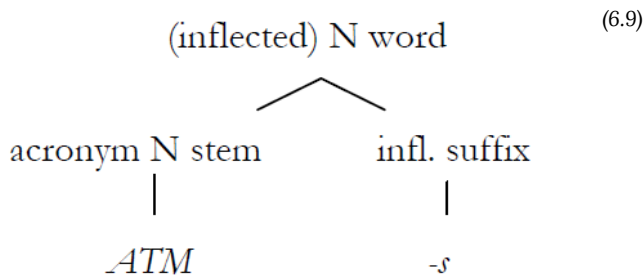
Figure 6.6. Examples of paraphrases of complex words

## 6.5.2 Representation of structure

The internal structure of words may be represented schematically by means of a diagram. In linguistics, diagrams that represent grammatical structure have become known as **tree diagrams**, although they in fact suggest an upside down, or inverted “tree”, with branches that grow downwards rather than upwards. By analogy with actual trees, tree diagrams have **branches**, straight lines that link units at successive levels of analysis, and **nodes**, the points at which the branching take place. Each node of the diagram bears a label, which clearly identifies the relevant unit for the intended analysis. Labelled tree diagrams are commonly used in morphology and in syntax, and their purpose is to enable us to visually grasp the linguistic structure of words, phrases and sentences in terms of their linear and hierarchical organisation.

When drawing a morphological tree diagram, we can work **bottom-up**, starting at the bottom of the tree, labelling each morpheme in each word, and work our way upwards. Or, we can work **top-down**, starting with the word as a whole and breaking it down into its constitutive morphemes. In either case, we must bear in mind that word analysis obeys the two constraints stated in section 4.2 above: the analysis reflects the hierarchical step-by-step process of word formation, and must build well-formed words at each stage of word formation.

Here are the complete diagrams for three words, the nouns *ATMs* and *piano-players*, and the verb *emailed*. For these diagrams, we chose, arbitrarily, to use the abbreviations *der.* and *infl.* for *derivational* and *inflectional*, and an arrow to indicate conversion. Other conventions can be used in diagrams, so long as their meaning is made perfectly clear.



The diagrams (6.9)–(6.11) give us all the information that we need in order to understand the internal structure of the words, in what could be called the words’ formation “history”. The information in brackets is in fact redundant, and is shown here just for clarity. There is no need to repeat, for example as in (6.10), that *player* is a derived stem: its suffix is already specified as derivational. The diagrams also show that inflection applies last, in all three word formations. All words in the examples are therefore inflected words, regardless of other processes in their formation.

#### Activity 6.14

Draw labelled tree diagrams for the underlined words in this sentence:

*Janice SMSed that her laptop refuses to restart properly.*

### 4.5.3 A note on spelling and morphological analysis

Written representations of language add an additional level of arbitrariness to it. We have also insisted that linguistics is concerned primarily with **spoken language**, rather than written/printed forms of it. “Creative spellers” do manage to get their written messages through, if their original spoken form can be recovered from the written/printed material. One example is the following letter written by a child to Santa Claus, where the intended meaning is clear despite the unexpected spelling:

(6.12) *I want a bored game.*

While the observation that spelling is a secondary representation

of language remains true, it is also true that **spelling** is not entirely irrelevant to linguistic analysis. Being conservative by nature, not least because it reinforces the dominance of the sense of sight over hearing, spelling preserves the visual coherence of morphologically-related words that may have lost their family resemblance in speech. Speech-faithful spellings like the following can be easily read, and might be advocated by spelling reformists:

(6.13) *ilektrik ilektrishan ilektrisiti ilektrikal*

The counterargument to such reformation is that the alternative spellings *ilektrik*, *ilektrish* and *ilektris* would fail to represent the unity of the morpheme *electric*, found in the conventional spelling of all four words.

## Food for thought

Let's face it,

English is a crazy language. There is no egg in the eggplant  
No ham in the hamburger

And neither pine nor apple in the pineapple. English muffins were  
not invented in England French fries were not invented in France.

We sometimes take English for granted

But if we examine its paradoxes we find that Quicksand takes you  
down slowly

Boxing rings are square

And a guinea pig is neither from Guinea nor is it a pig.

If writers write, how come fingers don't fing? If the plural of tooth  
is teeth

Shouldn't the plural of phone booth be phone beeth? If the  
teacher taught,

Why didn't the preacher praught?

If a vegetarian eats vegetables

What the heck does a humanitarian eat!? Why do people recite at  
a play

Yet play at a recital? Park on driveways and Drive on parkways?  
How can the weather be as hot as hell on one day And as cold as  
hell on another?

You have to marvel at the unique lunacy  
Of a language where a house can burn up as It burns down,  
In which you fill in a form By filling it out  
And a bell is only heard once it goes!  
English was invented by people, not computers And it reflects the  
creativity of the human race (Which of course isn't a race at all).

That is why  
When the stars are out they are visible  
But when the lights are out they are invisible. And why it is that  
when I wind up my watch It starts  
But when I wind up this poem It ends.

Richard Lederer

## Further reading

Deterding, David H. and Poedjosoedarmo, Gloria R. (2001). Chapter 2. Morphology. In *The grammar of English. Morphology and syntax for English teachers in Southeast Asia*. Singapore: Prentice Hall, pp. 6-17.

Hudson, Grover (2000). Chapter 15. Six ways to get new words. In *Essential introductory linguistics*. Oxford: Blackwell, pp. 239-251.

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.



# Chapter 7: History of English

# History of English

## Chapter Preview

In this unit you will:

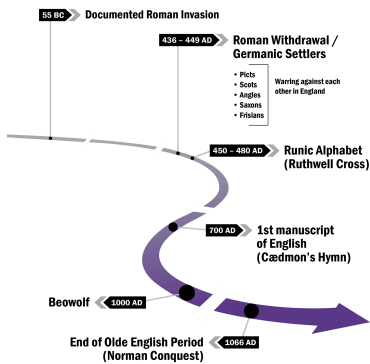
- Learn the history (and some very interesting stories) of English from Olde English, Middle English, and Early Modern English time periods
- Listen to Olde, Middle, and Early Modern English as it was spoken
- Practice speaking some of each time period

The history of our language is a colorful and exciting one. To begin, our language has not always sounded like it does today. Believe it or not, to today's speaker of English, our language may have sounded much more like German and Latin from its origins.

## Olde English

# A Little History

## Rome



### Old English timeline

When Rome was in the process of conquering the world, Englalond (England), was very much in their sights. Mostly because England has strong sea ports and trade routes already in place.

### Rome in England from 55 BC to 410 AD

By the 1st century AD, the Roman Empire governed all of the Italian peninsula, Romania, Switzerland, England, France, and most of the Mediterranean region as well as parts of North Africa.

### Latin

Language:

- Romans forced language change in many ways, but primarily through forcing merchants to do all their trade in Latin.
- Soon after infiltrating the commerce, they forced the use of



Roman Occupation 55 BC – 410 AD

Latin in government.

- Next, they forced the use of Latin in religion and education.

## The Germanic Invasion

### Rome Leaves England

When the Romans begin to leave England, other groups begin to invade. They were not as concerned with maintaining a language, so the use of Latin as the language of commerce, government, religion and education remained.

### Groups That Invade England Through 1066:



North—Irish/Gaelic, Scots,  
Picts

South/East—Angles, Saxons,  
Frisians

Throughout England—  
Scandinavian (Vikings)  
(See Map Left)

### **Why was Latin not the “common” language?**

Consider what you know already about the way the Romans forced the Latin language. Why, then, when the English went home, did they not speak Latin?

## **Scandinavian Influences**

The Scandinavians influences in the English language from about 700 to 1066 have

remained, for the most part, unchanged. They include:

### **Patronymics:**

The idea of knowing who you are based on your father's name. For example, if your father's name was Peter, you become Peter'sson or Peter'stadtter. So, Leonard Hofstadter (of The Big Bang fame) is actually an ancestor of a man whose name was Hof, who had a daughter!

### **Personal Pronouns:**

The Scandinavians brought with them the idea of third person plural pronouns.

The words include: they, them, their, and themselves. Can you imagine carrying on a conversation without those pronouns?

### **“sk”**

Most interesting is the sound signified by the Scandinavian letters [s k]. There is a story, that may or may not be true, the name for the garment 'skirt' or 'skyrt' was pronounced [sh ee r t a] by the Scandinavians. Now this was confusing for the English who also had a word 'scurt' which was pronounced [sh ee r t uh]. So, to reduce the confusion, the English began pronouncing the 'sc' as

we would pronounce the 'sk' today and 'scyrte' became 'skirt' and 'skyrtæ' became 'shirt'.

Interesting note: in many Norwegian countries today the 'sk' sound is pronounced [sh].

**Place Names:**

Many places in England were named by the Scandinavians. They used suffixes, such as: by, thwaite, and dale.

*Dal(e)*: to signify a valley



Yorkshire  
Dales

*Thwaite*: to signify a place in or near a meadow



Buttercup  
meadow

By: to signify to who owned the property



Magna Carta  
1225

## Ruthwell Cross (refer to Roman Occupation map above)

The story of the Ruthwell Cross and the story written on it is very important to our language.

The Ruthwell Cross gives us one of the first poems written in the Runic alphabet. This cross tells the story of Christ and his walk to Golgotha. What is most interesting is that this poem is told from the perspective of the cross.

The title of the poem is “The Dream of the Rood.” Rood in Olde English translates to wood.

“Krist wæs on rodi. Hwethræ ther fusæ ferran kwomu æththilæ til anum.”



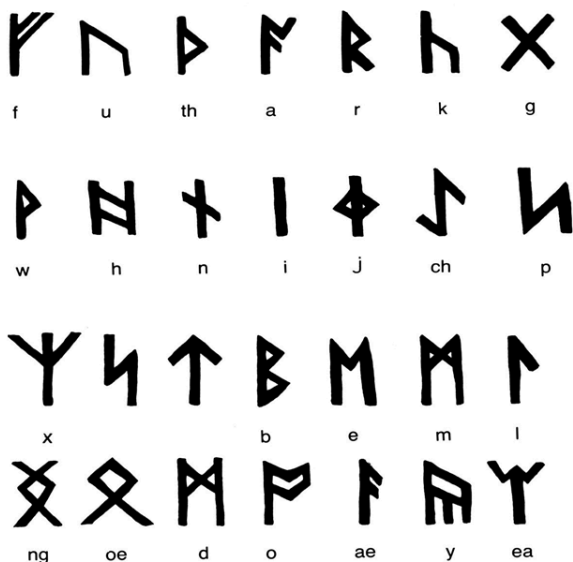
CC BY:  
*Daniel  
O'Donnell  
(July 25,  
2012)*

## Runes

It is not known exactly why the Runes have faded from existence. What is important, is that the pagan peoples who lived in the northern part of England, what is now Scotland, used this alphabet and now do not.

As English speakers and writers, have you considered what it might be like if English people had continued using this alphabet instead of the Roman alphabet?



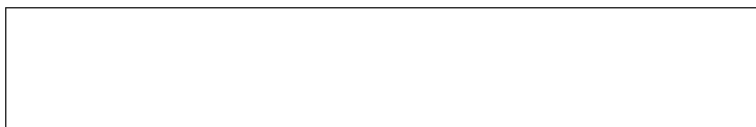


## Caedmon's Hymn

The story of Caedmon's Hymn is one of a man who wanted to be a priest. In the days of the Olde English time period, it was expected that priests would sing and write their own music. Caedmon was not talented in either of these (or so the story goes). He was a simple cow/sheep herder. His lifelong dream was to live among the priests at the abbey.

However, one night, while Caedmon slept, he had a dream that he wrote a song. He quickly shared this with the leader of the abbey where he lived.

His hymn is the oldest poem written in English. It is dated to approximately 658 and 680.





One or more interactive elements has been excluded from this version of the text. You can view them online

here: <https://pressbooks.utrgv.edu/introductiontolanguagestudies/?p=67#video-67-1>

**Musical Composition of the West Saxon version of Caedmon's Hymn, lyrics by Caedmon; music and performance by Clay Paramore, with Laura Aaron on piano. CC0.**

Northumbrian	West Saxon
Nū scylun hergan hefaenrīcaes Uard, metudæs maecti end his mōdgidanc, uerc Uuldurfadur, suē hē uundra gihwaes, ēci dryctin ōr āstelidæ hē ærist scōp aelda barnum heben til hrōfe, hāleg scepen. Thā middungeard moncynnæs Uard, eci Dryctin, æfter tiadæ firum foldu, Frēa allmectig	Nu sculon herigean heofonrices weard, meotodes meahte and his modgeþanc weorc wuldorfæder, swa he wundra gehwæs ece drihten, or onstealde. He ærest sceop eorðan bearnum heofon to hrofe, halig scyppend; þa middangeard moncynnes weard ece drihten, æfter teode firum foldan, frea ælmihtig
English translation (Carol Russell)	
Now we must praise the Guardian of heaven, The mighty Creator, and his creation The work of the Glorious Father, just as The mighty Lord began to establish his wonders. He, the Holy Creator, first fashioned the Earth for His children With heaven as a roof. Then mankind's Guardian, the almighty Lord, Afterwards adorned the earth with men.	

## Olde English Translation

Remember, in this time period, the vowels were nearly opposite of what we have today. Also, some of the sounds are different as well.

If you pay close attention to switching vowels and the information you know already from Caedmon's Hymn, you should be able to translate this poem as well.

**Original – This text is presented in the standardised West Saxon literary dialect.**

- [1] Fæder ure þu þe eart on heofonum,
- [2] Si þin nama gehalgod.
- [3] To becume þin rice,
- [4] gewurþe ðin willa, on eorðan swa swa on heofonum.
- [5] urne gedæghwamlican hlaf syle us todæg,
- [6] and forgyf us ure gyltas, swa swa we forgyfað urum
- [7] and ne gelæd þu us on costnunge, ac alys us of yfele. soþlice.

## Olde English Words

<b>Nature</b>	<b>Animals</b>	<b>Concepts</b>	<b>People</b>
æcer - field	æġ - eel	áð - oath	beorn - warrior
bæst - bast	bár - boar	borg - pledge	bydel - beadle
béam - tree	bucc - buck	céap - price	ceorl - churl
beorg - hill	bulluc - bullock	coss - kiss	cniht - boy
blóstm - blossom	earn - eagle	cræft - skill, strength	cyning - king
bóg - bough	eoh - horse	cwealm - death	dweorg - dwarf
bolt - bolt	eolh - elk	dóm - doom	eorl - nobleman
bróm - broom (the plant)	fearh - pig, boar	dream - joy, revelry	gást - spirit
clam - mud	fisc - fish	fæðm - embrace	hæft - captive
clút - patch	forsc - frog	fléam - flight	hwelp - whelp
cnoll - knoll	fox - fox	gang - going	m?g - kinsman
codd - cod, husk	géac - cuckoo	gielp - boasting	þegn - thane
cropp - sprout	h?ring - herring	hlæst - burden	þeof - thief
forst - frost	hengest - horse	hréam - cry, shout,uproar	wealh - foreigner
hægl - hail	hund - dog	torn - grief	wer - man
hærfest - autumn	hwæl - whale	þanc - thought	
healm - haul	mearh - horse	wæstm - growth	
hláf - loaf	seolh - seal (animal)		
horh - dirt	swertling - titlark		
hrim - rime	wulf - wolf		
hýdels - hiding place, cave			

<b>Nature</b>	<b>Animals</b>	<b>Concepts</b>	<b>People</b>
mæst - mast			
mór - moor			
múð - mouth			
regn - rain			
sealh - willow			
slóh - slough, mire			
stán - stone			
storm - storm			
stréam - stream			
swamm - swim			

# Order of Events

January	Edward	English
January-October	Harold	Anglo Saxon
October-December	Edgar	Hungarian
December 23rd	William	French

Edward is King until his death in January

- True “Anglo Saxon”
- Childless – uh oh- no heir apparent.

## The Cast



CCBY: Aidan Hart 3.0

**Edward  
The Confessor**



CCBY-SA 4.0

**Harold of England  
“Elected” as King**



CC0

**Edgar of Hungary**



CC0

**William of Normandy  
Claims to be successor**



CCBY-SA 2.0

**Harald of Norway  
Claims to be successor**



**Tostig, Harold's  
brother**

## The Conflict

There's the rub! King Edward has no children, but he does have a brother-in-law, Harold. Harold with an 'o'. So, he is 'elected' king. Elected by the Court, not the people.

The conflict continues as there are others who claim the throne, too.

- First, William of Normandy, who has some distant claim, that may not be strong
- Also, Edgar of Hungary, has a mother who is certain he has a claim to the throne
- Next in line is Harald of Norway also claims to be successor, as well

## The Plot Thickens

From April to September forces gather against each other

Harold had a brother (Tostig). He was estranged and lived in France. He sided with Harald and fought battles with him.

Harold had the support of the Normans. Remember, they are French.

William began gathering the Germanic peoples from villages all across England. By September he has a very strong army.

He has the support of nearly all of England's commoners. He cannot seem to conquer London.

NOTE: London is the seat of Government, Education, Religion. It has the busiest ports, and had the most inhabitants of all major cities.

## Through the months of September-December

Harold defeated and killed Harald and Tostig. William, however, loses London, but the rest of England submits to his power.

And then, Harold dies in battle—William looks like he might have England.

## But WAIT!

When Harold dies, England has no ruler.

Parliament names a boy of 14, who is a distant relative of Harold, king. His name is Edgar. He is Hungarian!



... And Then!



Detail from  
Bayeux  
Tapestry  
showing  
Bishop Odo  
rallying  
William the  
Conqueror's  
troops at the  
Battle of  
Hastings.

William defeats the Norman stronghold in London. He is crowned King on Dec. 23, 1066. There you have it...four rulers in one year. Two **Anglo Saxon**, one Hungarian, one French. This year, 1066, began for England two things. The first, the concept of pride of nationality. William helped with that by gathering the villagers together. The second was the idea that the aristocracy would befriend the French, or hate them. Depending on who was King or Queen, would depend on the relationship between England and France. The English language took (borrowed) words from the French, and the English aristocracy copied fashion, cuisine etc.

# Middle English

## Middle English Timeline



An interactive H5P element has been excluded from this version of the text. You can view it online here:

<https://pressbooks.utrgv.edu/introductiontolanguagestudies/?p=67#h5p-1>

## Changes begin!



*Flag of  
France  
(XII-XIII)*



*Saint George's cross (a red cross on a white background), used as the Flag of England, the Italian city of Genoa and various other places.*

It is during this time English leaders, who are sometimes French and sometimes English, either accept French influences or resist them.

One of the biggest changes in English when we move away from Old English to Middle English is that we begin our love/hate relationship with the French.

The earliest date of surviving written texts in Middle English dates to approximately 1150.

Much is happening in England during this time period. Oxford is founded as is Cambridge University.

By the late 1200's, English becomes the most common language spoken in England. By the 1300's the Great Vowel Shift begins changing English into something more akin to what our ears are accustomed to hearing.

The Great Vowel Shift occurs over several generations and about 100 or so years.

## The Great Vowel Shift

A VERY basic way to remember is that all short vowels became long and all long vowels became short

Dracula's law: 3rd shift=Blood, 2nd shift=good, 1st shift=food

“Blood is good food.” In Olde English it was “Blod is gud foud.

**Is this another “Great Vowel Shift”?**

Today in the Northern Great Lakes region of the US there is a vowel shift going on! “Bus” sounds to many Midwesterners like “Boss”

## Chaucer 1342-1400: The Canterbury Tales



Geoffrey  
Chaucer  
(c.1345-1400).  
English poet.  
Coloured full  
figure  
portrait.

Whan that Aprill, with his shoures soote

The droghte of March hath perced to the roote  
 And bathed every veyne in swich licour,  
 Of which vertu engendred is the flour;  
 5     Whan Zephirus eek with his sweete breeth  
       Inspired hath in every holt and heeth The tendre  
 croppes, and the yonge sonne  
       Hath in the Ram his halfe cours yronne,  
 And smale foweles maken melodye,  
 10    That slepen al the nyght with open eye-  
 (So priketh hem Nature in hir corages);  
       Thanne longen folk to goon on pilgrimages  
 And palmeres for to seken straunge strondes  
 To ferne halwes, kowthe in sondry londes;  
 15    And specially from every shires ende  
       Of Engelond, to Caunterbury they wende,  
       The hooly blisful martir for to seke  
       That hem hath holpen, whan that they were seeke.

This is an excerpt of the Prologue to the *Canterbury Tales*, 1386, written by Geoffrey Chaucer. It is a story of a group of people traveling together to visit religious sites in Canterbury, England.

Chaucer's English, although not entirely like Modern English, has more similarities than Olde English. As you can see (left), there are many more words that look and sound like modern English.

Notice, too, that the Great Vowel Shift has not been completed at this point.



One or more interactive elements has been excluded from this version of the text. You can view them online here: <https://pressbooks.utrgv.edu/introductiontolanguagestudies/?p=67#video-67-2>

CC BY, Ancient Literature Dude, The Canterbury Tales General Prologue, complete reading (Middle English, with timestamps)

## Caxton's Press

Caxton was an inventor, merchant, and ultimately a printer. In his travels as a merchant, he was intrigued with a printing press he saw in Cologne, France. He stayed there until he learned how to print pages.

He traveled back to England, where he built his own printing press and set up shop in Westminster Abbey.

The first book to have been printed on this press is Chaucer's *The Canterbury Tales*.

He is also credited for printing the first Bible verses on his press, as well.



William Caxton showing specimens of his printing to King Edward IV and his Queen. Published in *The Graphic* in 1877 referring to The Caxton Celebration. The Caxton Celebration, commemorating the 400th anniversary of the first printed book in England, took place in London in the summer of 1877

## Loan Words

Between 1300-1400

- The English language acquired 10,000+ loanwords
- 75% still used today



Pronunciation is changed due to outside influences – namely loanwords

- 31% of the loanwords are French
- Remaining – Latin, Scandinavian, Other

In the 1300's English acquires many more words that are added to the lexicon. Among the most used words today nearly 50% come from French or Latin. Although we credit many words from Scandinavian influences (remember they gave us place names and pronouns!)

### Exercises

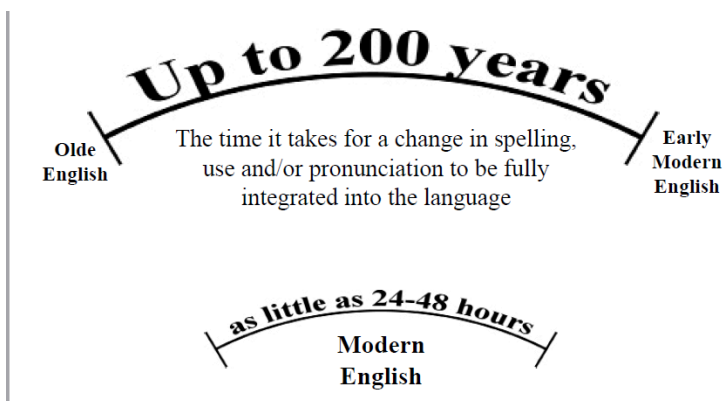
Using the list below, try to choose words in that category that you believe are NOT borrowed from another language.

- Government:
- Religion:
- Law:
- Military:
- Fashion:
- Food:
- Society:
- Art:
- Architecture:
- Literature:

- Medicine:

## Orthographic Lag

Orthographic Lag is the time it takes for a word, a pronunciation, or phrase to be added to the language and to become a permanent part of the lexicon.



### **...THINK ABOUT THIS:**

Why would the orthographic lag be reduced to such a short amount of time?

## Pronunciation Guide

In order to pronounce Middle English, speakers need only remember that vowels are not always pronounced the same as they are today.

One thing that will help you pronounce Middle English is to remember to drop the jaw and fully open the mouth when pronouncing vowels.

Use the following chart and practice on your own. In no time, you will be able to read *The Canterbury Tales* in its original version!

---

<b>VOWEL SOUNDS</b>	a = ah	e = ay	i / y = ee	o = oh	u = ôô
---------------------	--------	--------	------------	--------	--------

---

Symbol	As In:	Spelling	Examples:	
ă = ah	father	a, aa, er	fader, name, that, ferther, clerk, sterte	a
æ	hat, pass	a, ae	has, begge, great, heeth, baebe	
ay / ey	mate, day	e, ee	grene, sweete, me, be, she, fredom	e
ah-ee *	aisle	ai, ay, ei, ey	day, lai feith, veyne	
i *	my	ai, ay, ei, ey	fair, may, feith	
ar	there	e, ea	bere, sea	
a	sofa	first or final e	sonne, chivalrie, cause, yonge, name	
e	met	e, ee	ryden, hem, end, gentil	
ee-oo	few	eu, ew, u	new, reule	
ee	machine	i, y	passioun, shires, blithe, nyce, wif, ryde, icumen	i
i	bit	i, y	yloved, list, nyste, skille	
au	cloth	o, oo	lore, goon, ofte, hooly	
o = oh	sow, note	o, oo, ou, ow	bote, good, roote, thought, knowe	o
oo	root	u, ou, ow	juggen, resoun, flour, fowles, hous, vertu	u
o	book, full	o, oo, u	ful, nonne, love	
ah-oo	house, how	au, aw	cause, drawe	
aw-uh	Paul, awl	ou, ow	knownen, soule	
oy	boy	oi, oy	joye, point, coy	

# Early Modern English

## The Beginning of Modern English

### The Age of Reason Begins

By the time the Middle Ages pass there is a transformation of thinking. People are free, now, to pursue personal enlightenment and knowledge.

During the early 1600s, English is now, not only spoken in England, but is transplanted to North America with the founding of Jamestown.

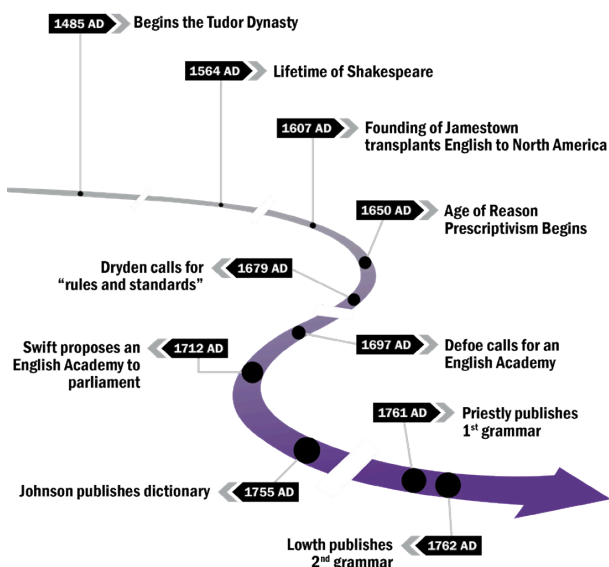
This also pushes the Age of Reason into the colonies. As man felt the call to find his own path in regards to religion, many felt they must leave England to do so.

It was during this time that leaders in circles of intellect and writing begin calling for changes in how the English view their language. Prescriptivism is born.



*Jamestown:  
the first  
permanent  
English  
speaking  
community  
in America.*

# Early Modern English Timeline



*Rules,  
Dictionary,  
Grammar: A  
Timeline*

## The Age of Reason: The Early Prescriptivists

### 1564-1616 Lifetime of Shakespeare

Although not always considered one of the founding Prescriptivists, Shakespeare did influence the language of those early Prescriptivists. He made up words, definitions, etc., which certainly may have helped those who were on the mission to set and refine the English language.

# 1611 King James Bible

Some claim that the King James Bible was printed on Caxton's Press. It is possible.

One interesting note is that Samuel Johnson's did use King James' Bible as a source for his definitions in his early editions of his dictionary.



Picture by  
Jeremylinvip:  
CCBY-SA 3.0

## 1679 Dryden wants “rules and standards”

John Dryden was most noted as being a straight forward poet and writer. Many claim that it is Dryden who established the rule that English sentences should not end in prepositions. He believed that standards for English would enhance not diminish its function.

## 1697 Defoe proposes an English Academy

Daniel Defoe is most famous for his novel, *Robinson Crusoe*. He was a prolific pamphleteer. He was imprisoned for a libel case. Yet he wrote in favor of creating an Academy in which English be established as the formal language of the unified England.

## 1712 Swift Proposes an English Academy to parliament

Jonathan Swift also wanted an academy and proposes the idea to Parliament. He believed in the power of the people and wanted a country where national pride was foremost.



# 1755 Johnson publishes dictionary

Title Page of  
Johnson's  
Dictionary

A  
**D I C T I O N A R Y**  
OF THE  
**E N G L I S H   L A N G U A G E :**  
IN WHICH  
The WORDS are deduced from their ORIGINALS,  
AND  
ILLUSTRATED in their DIFFERENT SIGNIFICATIONS  
BY  
EXAMPLES from the best WRITERS.  
TO WHICH ARE PREFIXED,  
A HISTORY of the LANGUAGE,  
AND  
AN ENGLISH GRAMMAR.  
By SAMUEL JOHNSON, A.M.  
IN TWO VOLUMES.  
VOL. I.  
THE SECOND EDITION.

*Cum vobis minimis cunctis finet bonelli:  
Anabit quocunque parum glimosa habebant,  
Et sine pondere erunt, et bonate indigne ferentur.  
Verba movere loco; quavis inrita recedant,  
Et veritate alius intra parcella Vellat:  
Oculosque dicit populo bonis erant, etque  
Proferret in lucem speciosa vocabula rerum,  
Que pille memorata Camillus atque Cestrigis,  
Nunc suis infernis pennis et detersa vetula.*

Hoz.

LONDON,  
Printed by W. STRAHAN,  
For J. and P. Knapton; T. and T. Longman; C. Hitch and L. Hawes;  
A. Millar; and R. and J. Dodsley.  
MDCCCLV.

Samuel Johnson writes the first English dictionary. He was under contract to write his dictionary in 1746. It was published in 1755. He was known to state that what took the French Academy nearly 40

years to do for their language was accomplished by one man in only 9 years.

1761 Priestley publishes first grammar book

THE  
R U D I M E N T S  
O F  
ENGLISH GRAMMAR;  
ADAPTED TO THE  
U S E of S C H O O L S.  
WITH  
O B S E R V A T I O N S on STYLE.  
By JOSEPH PRIESTLEY.

Title Page of  
Priestley's  
Book

L O N D O N :  
Printed for R. GRIFFITHS, in the Strand.  
M.DCC.LXI.

Joseph Priestley is the first to write a book on English grammar in, *Rudiments of English Grammar*.

## 1762 Lowth publishes second grammar book

Just a year later, Robert Lowth publishes his grammar book, *A Short Introduction to English Grammar, With Critical Notes*.

It is from these men we have what is called Prescriptivism, today. These men wanted to prescribe for its speakers what the English language ought to be and to refine the language. They sought to keep language in a fixed and permanent state.

## Prescriptivism

### What is Prescriptivism?

Prescriptivism is prescribing language.

It is how a group of 6 men wanted to 'fix' the English language in a permanent way. To put language in a box and keep it from changing.

What these six guys stood for was order and regulation.

What they did not sanction among many others was clipped words, contractions, and slang.

Think of the language we speak today. Can you imagine speaking and never using a contraction or a word that has been clipped?

And, oh my, the slang we use today! I imagine the 6 guys would turn over in their graves to hear the language we use today.

### **End Results**

The end result of prescriptivism left people thinking that language must be fixed. Prescriptivism did not allow for changes in the language either in pronunciation or spelling. Prescriptivism sought to logically order language using grammar rules and regulations, which we have adopted and maintained ever since.

## **The Growth of our Lexicon**

### **Derivation**

These kinds of words are very common and have been useful in creating ways of speaking and being understood.

You begin with a root word like 'lock'. By adding a prefix or a suffix a new word can be created.

This is especially helpful in understanding if some can or cannot be done, in this case, locked.

*Examples*

**How to make:**

root word + suffix/prefix = new word

UN+LOCK+ABLE = Unlockable

### **Compounding**

These kinds of words are very common today. Compounding begins with two words that do not have anything in common.

In this case a neck and a lace. By combining the two, we get a new word 'necklace'.

Another favorite is dug and out. Yet, when we compound them, we get a new noun which we all know as dugout.

#### *Examples*

#### **How to make:**

Word A + Word B = Word AB

Neck + Lace = Necklace

#### *Attribution*

This chapter has been adapted from Linguistics for Teachers of English by Carol Russell under a CC BY-NC-SA 4.0 license.

# Chapter 8: Dialects

Use the link here to access *Studying Language Variation*, chapter 8 “Sociolinguistics” via the UTRGV Library.

<https://www-taylorfrancis-com.ezhost.utrgv.edu/chapters/mono/10.4324/9780429444623-8/studying-language-variation-kristin-denham-anne-lobeck?context=ubx&refId=a47900ec-9f9d-4049-af4f-21e14ada5372>

# Technical notation and terms

Below is a summary of the notation conventions used in this book, followed by a table of technical terms. The table shows the different forms that each term can take, to help you select the appropriate one for your purposes.

## Summary of notation conventions

### X\_Z

Represents the distribution frame of a linguistic unit. The blank represents the relevant unit. X and Z represent the linguistic units immediately before and after the relevant unit, e.g.

Det\_\_N represents the distribution frame of Adj

Det represents the distribution frame of N

### { }

Curly brackets are used in morphology to enclose morphemes, e.g.

the word *unhappiness* comprises the morphemes {*un-*}, {*happy*}, and {*-ness*}

the word *challenging* comprises the morphemes {*challenge*} and {*-ing*}

Curly brackets (or braces) are also used in syntax, in phrase structure (PS) rules, to denote alternative constituents, e.g.

$$\text{NP} \rightarrow \left\{ \begin{array}{c} \text{Det (Adj)* N} \\ \text{Pr} \end{array} \right\}$$

[ ]

Phonetic symbols appear within square brackets, e.g. [f], [b], [m]

Phonetic transcriptions appear within square brackets.

Allophones appear within square brackets.

/ /

Phonemes appear within slashes, e.g. /f/, /b/, /m/ Phonemic representations appear within slashes.

→

Used in PS rules to mean 'expands into', 'is constituted by', 'is rewritten as'

()

Used in PS rules to denote optional constituents

\*

An asterisk before an example indicates that the use of language in the example is ungrammatical (syntactically unacceptable or non-occurring)

An asterisk after a linguistic unit indicates one or more of that unit

#

before an example, indicates that the use of language in the example is semantically odd or unacceptable.

Technical terms and how to use them  
appropriately



<b>Noun use</b>	<b>Verb use</b>	<b>Adjective use</b>	<b>Adverb use</b>	<b>Agent use</b>
active	-	active	actively	-
adjective	adjectivise	adjectival	adjectivally	-
adjunct	adjoin	adjunct	-	-
adverb	-	adverbial	adverbially	-
affix affixation	affix	affixed (word)	-	-
agreement	agree	-	-	-
ambiguity	-	ambiguous	ambiguously	-
anaphora	anaphorise	anaphoric	anaphorically	-
articulation	articulate	articulated	-	articulator
compound	compound	compounded	-	-

conjunction	conjoin	conjunctional	conjunctively	-
conversion	convert	converted	-	-
coordination	coordinate	coordinate(d)	coordinately	coordinator
derivation	derive	derived (word) derivational (affix)	derivationally	-
diagram	diagram	diagrammatic	diagrammatically	-
grammar	grammaticise	grammatical	grammatically	grammarian

<b>Noun use</b>	<b>Verb use</b>	<b>Adjective use</b>	<b>Adverb use</b>	<b>Agent use</b>
head	head	headed	–	–
inflection	inflect	inflected (word) inflectional (affix)	inflectionally	–
lexicon	lexicalise	lexical	lexically	lexicologist
linguistics	–	linguistic	linguistically	linguist
morphology	–	morphological	morphologically	morphologist
noun	nominalise	nominal	nominally	–
object	–	object	–	–
passive	passivise	passive	passively	–
phonetics	–	phonetic	phonetically	phonetician
phonology phoneme	–	phonological phonemic	phonologically phonemically	phonologist
phrase	phrase	phrasal	–	–
plural	pluralize	plural	–	–
preposition	–	prepositional	prepositionally	–
pronoun	pronominalise	pronominal	pronominally	–
science	–	scientific	scientifically	scientist
semantics	–	semantic	semantically	semantician semanticist
singular	–	singular	–	–
subject	–	subject	–	–
subordination	subordinate	subordinate(d)	subordinately	subordinator
syntax	–	syntactic	syntactically	syntactician
verb	verbalise	verbal	verbally	–

### *Attribution*

This chapter has been modified and adapted from The Language of Language. A Linguistics Course for Starters under a CC BY 4.0 license. All modifications are those of Régine Pellicer and are not reflective of the original authors.