

Chapter 1

Introduction

1.1

WHY STUDY WORDS?

Imagine a life without words! Trappist monks opt for it. But most of us would not give up words for anything. Every day we utter thousands and thousands of words. Communicating our joys, fears, opinions, fantasies, wishes, requests, demands, feelings—and the occasional threat or insult—is a very important aspect of being human. The air is always thick with our verbal emissions. There are so many things we want to tell the world. Some of them are important, some of them are not. But we talk anyway—even when we know that what we are saying is totally unimportant. We love chitchat and find silent encounters awkward, or even oppressive. A life without words would be a horrendous privation.

It is a cliché to say that words and language are probably humankind's most valuable single possession. It is language that sets us apart from our biologically close relatives, the great primates. (I would imagine that many a chimp or gorilla would give an arm and a leg for a few words—but we will probably never know because they cannot tell us.) Yet, surprisingly, most of us take words (and more generally language) for granted. We cannot discuss words with anything like the competence with which we can discuss fashion, films or football.

We should not take words for granted. They are too important. This book is intended to make explicit some of the things that we know subconsciously about words. It is a linguistic introduction to the nature and structure of English words. It addresses the question 'what sorts of things do people need to know about English words in order to use them in speech?' It is intended to increase the degree of sophistication with which you think about words. It is designed to give you a theoretical grasp of English word-formation, the sources of English vocabulary and the way in which we store and retrieve words from the mind.

I hope a desirable side effect of working through *English Words* will be the enrichment of your vocabulary. This book will help to increase, in a very practical way, your awareness of the relationship between words. You will be equipped with the tools you need to work out the meanings of unfamiliar words and to see in a new light the underlying structural patterns in many familiar words which you have not previously stopped to think about analytically.

For the student of language, words are a very rewarding object of study. An understanding of the nature of words provides us with a key that opens the door to an understanding of important aspects of the nature of language in general. Words give us a panoramic view of the entire field of linguistics because they impinge on every aspect of language structure. This book stresses the ramifications of the fact that words are complex and multi-faceted entities whose structure and use interacts with the other modules of the grammar

2 ENGLISH WORDS

such as PHONOLOGY, the study of how sounds are used to represent words in speech, SYNTAX, the study of sentence structure, and SEMANTICS, the study of meaning in language.

In order to use even a very simple word, such as *frog*, we need to access various types of information from the word-store which we all carry around with us in the MENTAL LEXICON or DICTIONARY that is tucked away in the mind. We need to know:

[1.1]

- (i) its shape, i.e. its PHONOLOGICAL REPRESENTATION/*frg*/ which enables us to pronounce it, and its ORTHOGRAPHIC REPRESENTATION *frog*, if we are literate and know how to spell it (see the Key to symbols used on page xix);
 - (ii) its grammatical properties, e.g. it is a noun and it is countable—so you can have one *frog* and two *frogs*;
 - (iii) its meaning.
-

But words tend not to wear their meaning on their sleeve. Normally, there is nothing about the form of words that would enable anyone to work out their meaning. Thus, the fact that *frog* refers to one of these simply has to be listed in the lexicon and committed to memory by brute force. For the relationship between a LINGUISTIC SIGN like this word and its meaning is ARBITRARY. Other languages use different words to refer to this small tailless amphibian. In French it is called (*la*) *grenouille*. In Malay they call it *katak* and in Swahili *chura*. None of these words is more suited than the others to the job of referring to this small reptile.

And of course, within a particular language, any particular pronunciation can be associated with any meaning. So long as speakers accept that sound-meaning association, they have a kosher word. For instance, *convenience* originally meant ‘suitability’ or ‘commodiousness’ but in the middle of the nineteenth century a new meaning of ‘toilet’ was assigned to it and people began to talk of ‘a public convenience’. In the early 1960s the word acquired the additional new meaning of ‘easy to use, designed for hassle-free use’ as in *convenience food*.

We are the masters. Words are our servants. We can make them mean whatever we want them to mean. Humpty Dumpty had all this worked out. The only thing missing from his analysis is the social dimension. Any arbitrary meaning assigned to a word needs to be accepted by the speech community which uses the language. Obviously, language would not be much use as a means of communication if each individual language user assigned a private meaning to each word which other users of the language did not recognise. Apart from that, it is instructive to listen in on the lesson on the nature of language that Humpty Dumpty gave to Alice (see overleaf).

Let us now consider one further example. All competent speakers of English know that you can add *-s* to a noun to indicate that it refers to more than one entity. So, you say *cat* when referring to one and *cats* if there is more than one. If you encountered in the blank in [1.2a] an unfamiliar word like *splet* (which I have just made up), you would automatically know from the context that it must have the plural form *splets* in this position since it is specified as plural by *all*. Further, you would know that the plural of *splet* must be *splets* (rather than *spletren* by analogy to *children* or *spleti* by analogy to *stimuli*). You know that the majority of nouns form their plural by adding the regular plural suffix or ending *-s*. You always add *-s* unless express instructions are given to do otherwise. There is no need to memorise separately the plural form of most nouns. All we need is to know the rule that says ‘add *-s* for plural’. So, without any hesitation, you suffix *-s* to obtain the plural form *splets* in [1.2b]:

Chapter 1

Introduction

1.1

WHY STUDY WORDS?

Imagine a life without words! Trappist monks opt for it. But most of us would not give up words for anything. Every day we utter thousands and thousands of words. Communicating our joys, fears, opinions, fantasies, wishes, requests, demands, feelings—and the occasional threat or insult—is a very important aspect of being human. The air is always thick with our verbal emissions. There are so many things we want to tell the world. Some of them are important, some of them are not. But we talk anyway—even when we know that what we are saying is totally unimportant. We love chitchat and find silent encounters awkward, or even oppressive. A life without words would be a horrendous privation.

It is a cliché to say that words and language are probably humankind's most valuable single possession. It is language that sets us apart from our biologically close relatives, the great primates. (I would imagine that many a chimp or gorilla would give an arm and a leg for a few words—but we will probably never know because they cannot tell us.) Yet, surprisingly, most of us take words (and more generally language) for granted. We cannot discuss words with anything like the competence with which we can discuss fashion, films or football.

We should not take words for granted. They are too important. This book is intended to make explicit some of the things that we know subconsciously about words. It is a linguistic introduction to the nature and structure of English words. It addresses the question 'what sorts of things do people need to know about English words in order to use them in speech?' It is intended to increase the degree of sophistication with which you think about words. It is designed to give you a theoretical grasp of English word-formation, the sources of English vocabulary and the way in which we store and retrieve words from the mind.

I hope a desirable side effect of working through *English Words* will be the enrichment of your vocabulary. This book will help to increase, in a very practical way, your awareness of the relationship between words. You will be equipped with the tools you need to work out the meanings of unfamiliar words and to see in a new light the underlying structural patterns in many familiar words which you have not previously stopped to think about analytically.

For the student of language, words are a very rewarding object of study. An understanding of the nature of words provides us with a key that opens the door to an understanding of important aspects of the nature of language in general. Words give us a panoramic view of the entire field of linguistics because they impinge on every aspect of language structure. This book stresses the ramifications of the fact that words are complex and multi-faceted entities whose structure and use interacts with the other modules of the grammar

The following chapter continues the discussion of the role of the lexicon. It attempts to answer questions like 'what is the lexicon for?' 'What items need to be listed in the dictionary?' 'What is the difference between idioms (like *to nail one's colours to the mast*) and syntactic phrases (like *to nail a notice to the door*)?' The next two chapters highlight the fact that the English word-store is vast and infinitely expandable. First, in Chapter 9 we consider the ways in which, using the internal resources of the language, speakers are able to produce an indefinitely large number of words. In Chapter 10 attention shifts to the expansion of English vocabulary through the importation of countless words from other languages. The story of imported words is in many ways also the story of the contacts that speakers of English have had with speakers of other languages over the centuries.

Most of the space in this book is devoted to an examination of the structure of English words. But the analysis of word-structure is seen not as an end in itself, but rather as a means to an end. And that end is to understand what it means to *know* a word. What sorts of information about words do you need to have in order to use them in communication? So the final chapter is devoted to the MENTAL LEXICON. It addresses the question, 'how is it that people are able to store a vast number of words in the mind and to retrieve the right one so fast in communication?' We will see that words are not piled in a muddle in the mind. Rather, the mental lexicon is very highly organised. This concluding chapter will also pull together the various strands developed in the earlier chapters.

I have already stressed the point that morphology is not a self-contained module of language. Any discussion of word-formation touches on other areas of linguistics, notably phonology and syntax, so I have provided a key to the list of pronunciation symbols at the beginning of the book. I have also included at the end a glossary of linguistic terms (many of them from other branches of linguistics) which might be unfamiliar. But still I may have missed out some terms. If you encounter any unfamiliar technical terms that are not explained in this book, I suggest that you consult a good dictionary of linguistics like Crystal (1991). Sometimes it is useful to present data using phonetic notation. A key to the phonetic symbols used is to be found on pp. xix–xx.

After this introductory chapter, all chapters contain exercises. Several of the analytical exercises require you to look up words and parts of words in a good dictionary like the *Oxford English Dictionary*. Access to such a dictionary is essential when you study this book. This is a practical way of learning about the structure of English words (and may also be a useful way of enriching your vocabulary).

Chapter 2

What is a word?

2.1 INTRODUCTION

Often we find it very difficult to give a clear and systematic account of everyday things, ideas, actions and events that surround us. We just take them for granted. We rarely need to state in an accurate and articulate manner what they are really like. For instance, we all know what a game is. Yet, as the philosopher Wittgenstein showed, we find it very difficult to state explicitly what the simple word *game* means.

The same is true of the term *word*. We use words all the time. We intuitively know what the words in our language are. Nevertheless most of us would be hard pushed to explain to anyone what kind of object a word is. If a couple of Martian explorers (with a rudimentary understanding of English) came off their space-ship and stopped you in the street to enquire what earthlings meant by the term WORD what would you tell them? I suspect you might be somewhat vague and evasive. Although you know very well what words are, you might find it difficult to express explicitly and succinctly what it is that you know about them.

The purpose of this chapter is to try to find an answer to the question: what is a word? It is not only Martian explorers curious about the way earthlings live who might want to know what words are. We too have an interest in understanding words because they play such an important role in our lives. As we saw in the last chapter, it is impossible to imagine human society without language. And equally, it is impossible to imagine a human language that has no words of any kind. It is impossible to understand the nature of language without gaining some understanding of the nature of words. So, in this chapter we will clarify what we mean when we use the term 'word'. This clarification is essential if our investigations are to make any headway for, as you will see presently, we mean quite a few very different things when we talk of words.

A standard definition of the word is found in a paper written in 1926 by the American linguist Leonard Bloomfield, one of the greatest linguists of the twentieth century. According to Bloomfield, 'a minimum free form is a word'. By this he meant that the word is the smallest meaningful linguistic unit that can be used on its own. It is a form that cannot be divided into any smaller units that can be used independently to convey meaning. For example *child* is a word. We cannot divide it up into smaller units that can convey meaning when they stand alone.

Contrast this with the word *childish* which can be analysed into *child-* and *-ish*. While the *child* bit of *childish* is meaningful when used on its own (and hence is a word), the same is not true of *-ish*. Although according to the *Oxford English Dictionary (OED)* *-ish* means something like 'having the (objectionable) qualities of' (as in *mannish*, *womanish*, *devilish*, *sheepish*, *apish* etc.), there is no way we can use it on its

the meaning of 'someone who does whatever the verb means'. Given the verb *tickoff*, a *ticker-off* must be a person who *ticks off*. Similarly, if you know what established words like *handful*, *cupful* and *spoonful* mean, you are also able to figure out the meanings of novel words like *fountain-penful* (as in *a fountain-penful of ink*) or *hovercraftful* (as in *hovercraftful after hovercraftful of English shoppers returned from Calais loaded down with cigarettes, cheese and plonk*). Virtually any noun denoting a container can have *-ful* added to it in order to indicate that it is 'full of something'.

To take another example, a number of words ending in *-ist*, many of which have come into use in recent years, refer to people who discriminate against, or hold negative views about, certain less powerful subgroups in society, e.g. *racist*, *sexist*. Anyone who knows what *racist* and *sexist* mean, given the right context should have no difficulty in understanding the nature of discrimination perpetrated by people who are described using the novel words *ageist*, *sizist* and *speechist*. *Ageism* is discrimination on grounds of (old) age—for instance, denying employment to people over the age of 60; *sizism* is discrimination (usually against fat people) on grounds of size and *speechism* is discrimination against people with speech impediments like stuttering.

Did you notice how I exploited your tacit knowledge of the fact that words ending in *-ist* and *-ism* complement each other? You were glad to accept *ageism*, *sizism* and *speechism* because you know that corresponding to an adjective ending in *-ist* there will normally be a noun ending in *-ism*. This is important. It shows that you know that certain word-forming bits go together—and others do not. I suspect that you would reject putative words like **agement*, **sizement* and **speechment*. (An asterisk is used conventionally to indicate that a form is disallowed.) In word-formation it is not a case of anything goes.

A challenging question which morphology addresses is, 'how do speakers know which non-occurring or non-established words are permissible and which ones are not?' Why are the words *fountainpenful*, *hovercraftful* and *speechist* allowed while **agement*, **speechment* and **sizement* are not?

Morphological theory provides a general theory of wordformation applicable to any language but, as mentioned earlier, this book focuses on word-formation in English. Its objective is to provide a description of English words designed to make explicit the various things speakers know, albeit in an unconscious manner, about English words. The emphasis will be on the description of English words rather than the elaboration of morphological theory. So, data and facts about English words are brought to the fore and the theoretical and methodological issues are kept in the background for the most part. The use of formal notation has also been kept to a minimum in order to keep the account simple.

1.2

OVERVIEW OF COMING CHAPTERS

At the very outset we need to establish the nature of the subject we are going to be examining. So, Chapter 2 discusses the nature of words. Then the next three chapters delve deep inside words and investigate their internal structure. In the process, traditional morphological concepts of structural linguistics are introduced and extensively exemplified.

Morphology is not a stand-alone module. After the introductory chapters, in Chapter 6 you are introduced to a theory where morphology is an integral part of the LEXICON or DICTIONARY. This chapter focuses on the interaction of phonology and morphology in word-formation.

Chapter 7 explores the relationship between words in speech and in writing. What is the relationship between saying words and writing them down? Is writing simply a mirror of speech—and an apparently distorting one in the case of English?

own. If some shouted to you in the street, 'Hey, are you *-ish?*' you might smile bemusedly and think to yourself, 'Isn't he weird!' In the next chapter we will take up the question of what to do with pieces of words that cannot be used meaningfully on their own. But for the moment we will focus exclusively on words.

2.2 WORDS ARE LIKE LIQUORICE ALLSORTS

When we talk of words we do not always mean exactly the same thing. Like liquorice allsorts, words come in all sorts of varieties. We will start our discussions by distinguishing the different senses in which we use the term 'word'.

2.2.1 Word-forms

Let us use the term WORD-FORM to describe the physical form which realises or represents a word in speech or writing. Consider the words in the following extract from T.S.Eliot's poem:

[2.1]

Half-past one,
The street-lamp sputtered,

The street-lamp muttered,
The street-lamp said, 'Regard that woman
Who hesitates towards you in the light of the door
Which opens on her like a grin...

(*'Rhapsody on a windy night'* in Eliot 1963)

In written English, words are easy to recognise. They are preceded by a space and followed by a space. Using this criterion, we can say that there are thirty-one words (i.e. word-forms) in the extract from 'Rhapsody'. We will call word-forms like these which we find in writing ORTHOGRAPHIC WORDS. If you look again at the extract, you might wonder if some of the hyphenated orthographic words are 'really' individual words. Many people would hyphenate *half-past* as Eliot does but not *street-lamp*. They would write *street lamp* as two separate words, with a space between them. What would you do?

The use of hyphens to indicate that something is a complex word containing more than one word-like unit is variable, largely depending on how transparent the compound nature of a word is. Shakespeare wrote *today* as *to-day* and *tomorrow* as *to-morrow*:

[2.2]

a. To-morrow, Caesar,
I shall be furnished to inform you rightly...
(*Antony and Cleopatra*, I, iv)

- b. O! that we now had here
 But ten thousand of those men in England
 That do not work to-day.
 (*Henry V*, IV, iii)
-

Hyphenating *to-day* and *to-morrow* is less common now, probably because most speakers are unaware of the compound nature of these words. *Today* comes from Old English *t dae* 'to+day' and *tomorrow* is from Middle English *to mor(e)we* (i.e. to (the) morrow) —*to-* can be traced back ultimately to a form that meant 'this' in Indo-European. Note in passing that three major periods are distinguished in the history of the English language: Old English (conventionally abbreviated as OE) was spoken c.450–1100; Middle English (conventionally abbreviated as ME) was spoken c.1100–1500 and Modern English from 1500 to the present.

Generally, the use of the hyphen in such words that are no longer seen as compounds is in decline. The hyphen tends to be mostly used in compounds that are regarded as fairly new words. Many well-established words that are transparently compounded, e.g. *schoolboy*, are normally written without a hyphen. Of course, judgements as to what is an established word vary greatly. There are few firm rules here. For instance, in the *OED* both *seaway* and *sea-way* are shown to be accepted ways of writing the word pronounced as /si:weɪ/. Similarly, the compilers of the *OED* show variation in the way they enter both hyphenated *first-rate* and *first rate* written as two words separated by a space.

Interestingly, hyphenation is also used creatively to indicate that an idea that would normally be expressed by a phrase is being treated as a single word for communicative purposes because it has crystallised in the writer's mind into a firm, single concept. Thus, for example, the expression *simple to serve* is normally a phrase, just like *easy to control*. But it can also be used as a hyphenated word as in *simple-to-serve recipe dishes* (*M&S Magazine* 1992:9). Similarly, on page 48 of the same magazine, the writer of an advertising feature uses the phrase 'fresh from the farm' as a hyphenated word in 'fresh-from-the-farm eggs'. But for creative hyphenation you are unlikely to find anything more striking than this:

[2.3]

On Pitcairn there is little evidence of the *what-we-have-wehold, no-surrender, the Queen's-picture-in-every-room* sort of attitude.

Simon Winchester in *The Guardian* magazine, 12 June 1993: 27; (italics added to highlight the compounds)

What we have established is that as a rule, orthographic words have a space on either side of them. But there are cases where this simple rule of thumb is not followed. There is a degree of flexibility in the way in which words are written down: being, or not being, separated by a space is in itself not a sure sign of word status. Some orthographic words which are uncontroversially written as one unit contain two words within them. They are compound words like *firstrate*, *seaway*, *wheelbarrow* and *teapot*. Furthermore, there are forms like *they're*, *hadn't* and *I'm* which are joined together in writing yet which are not compound words. When you scratch the skin, you see immediately that *they're*, *hadn't* and *I'm* are really versions of the pairs of words *they are*, *had not* and *I am*. Our theory needs to say something about awkward customers like these. Since the issues they raise are complex, we will postpone discussion of them until sections (4.3) and (8.3). Finally, there are words which are compounded (and maybe hyphenated as in [2.3]) as a one-off to crystallise a particular meaning.

ˈacrobat	aˈnnoying	caˈhoots
ˈkingfisher	deˈmolish	gaberˈdine
ˈpatriarchate	Chauˈcerian	hullabaˈloo

Main stress can fall on only one syllable in a word. The location of main stress is part of the make-up of a word and is not changed capriciously by individual speakers. You cannot decide to stress *hullabaloo* on the penultimate syllable on a Monday (*hullaˈbaloo*), on the antepenultimate syllable on a Tuesday (*huˈllabaloo*), on the initial syllable on a Wednesday (*ˈhullabaloo*) and on the final syllable for the rest of the week (*hullabaˈloo*).

However, in some cases, if we wish to contrast two related words, we can shift stress from its normal position to a new position. This can be seen in *ˈvendor* and *venˈdee* which normally are stressed on the first and second syllable respectively. But if the speaker wants to contrast these two words both words might be stressed on the final syllable as I heard an estate agent do in a radio interview.

[2.6]

It is *venˈdor*, not the *venˈdee* who pays that tax.

This example illustrates well the point that a word is allowed just one stress. Stress can be shifted from one syllable to another, but a word cannot have two main stresses. We could not have **ˈvenˈdor* and **ˈvenˈdee* where the two syllables received equal stress. Stress has to do with relative prominence. The syllable that receives main stress is somewhat more prominent than the rest, some of which may be unstressed or weakly stressed. By contrast, function words are normally unstressed. We can say *Nelly went to town* with no stress on *to* unless we wish to highlight *to* for contrastive purposes, e.g. *Nelly went to town and not far away from town*).

It is easy to see how stress can function as a valuable clue in determining whether two content words are a single compound word or two separate words. The nouns *street* and *lamp* are both stressed when they occur in isolation. But if they appear in the compound *ˈstreet-lamp*, only the first is stressed. The stress on *lamp* is suppressed.

Stress is not the only phonological clue. In addition to stress, there are rules regulating the positions in which various sounds may occur in a word and the combinations of sounds that are permissible. These rules are called PHONOTACTIC RULES. They can help us to know whether we are at the beginning, in the middle or at the end of a word. A phonological word must satisfy the requirements for words of the spoken language. For instance, while any vowel can begin a word, and most consonants can appear alone at the beginning of a word, the consonant [ŋ] is subject to certain restrictions. (This consonant is spelled *ng* as in *long* (see the Key to symbols used on p. xix). In English words [ŋ] is not allowed to occur initially although it can occur in other positions. Thus, [ŋ] is allowed internally and at the end of a word as in [l ŋ] *longing* and [l ɡe] *longer*. But you could not have an English word like *ngether*, *[ŋ ee] with [ŋ] as its first sound. However, in other languages this sound may be found word-initially as in the Chinese name *Nga* [ŋ a] and the Zimbabwean name *Nkomo* [ŋ komo].

There are also phonotactic restrictions on the combination of consonants in various positions in a word in the spoken language. As everyone knows, English spelling is not always a perfect mirror of pronunciation. So when considering words in the spoken language it is important to separate spelling from pronunciation (cf. Chapter 7). You know that *He is knock-kneed* is pronounced /hɪ ɪz nɒk ni:d/ and not **/he ɪs knk kni:d/*. A particular combination of letters can be associated with very different pronunciations in different words or

So far we have only considered orthographic words, i.e. recognisable physical written word-forms. Obviously, words as physical objects exist not only in writing, but also in speech. We will now briefly turn to word-forms in spoken language. We will refer to them as PHONOLOGICAL WORDS.

The challenge of word recognition arises in an even more obvious way when we consider speech. Words are not separated distinctly from each other. We do not leave a pause between words that could be equated to a space in writing. (If we did that, conversation would be painfully slow! Just try speaking to one of your friends today leaving a two-second gap between words. See how they react.) In normal speech words come out in a torrent. They overlap. Just as droplets of water cannot be seen flowing down a river, individual words do not stand out discretely in the flow of conversation. So they are much harder to isolate than words in writing. None the less, we are able to isolate them. If you heard an utterance like:

[2.4]

The cat slept in your bed.
/ekæt slept In : bed/

(Note: ‘ˈ’ shows that the following syllable is stressed; phonemic transcription is written between slant lines.)

you would be able to recognise the six phonological words that have been written in PHONEMIC TRANSCRIPTION (which shows the PHONEMES, i.e. the sounds that are used to distinguish the meanings of words) although what you hear is one continuous stream of sound. For purely practical reasons, throughout the book, unless otherwise stated, phonemic transcriptions and references to pronunciation will be based on RECEIVED PRONUNCIATION (RP), the prestige accent of standard British English—the variety popularly known as the Queen’s English or BBC English.

An intriguing question that linguists and psychologists have tried to answer is: how do people recognise words in speech? We will address this question in detail in section (11.2.1) below. For now let us simply assume that phonological words can be identified. Our present task will simply be to outline some of their key properties. To do this it will be useful to distinguish between two types of words: the so-called CONTENT WORDS and FUNCTION WORDS. Content words are the nouns, verbs, adjectives and adverbs which contain most of the REFERENTIAL (or COGNITIVE MEANING) of a sentence. This roughly means that they name individuals and predicate of them certain properties. They tell us, for instance, what happened or who did what to whom, and in what circumstances. An example will make the point clear. In the old days, when people sent telegrams, it was content words that were mainly (or exclusively) used. A proud parent could send a message like *Baby girl arrived yesterday* which contained two nouns, a verb and an adverb. Obviously, this is not a well-formed, grammatical sentence. But its meaning would be clear enough.

Function words are the rest—prepositions, pronouns, conjunctions, articles and so on. They have a predominantly grammatical role. A telegram containing only the words *She it and for us* would convey little idea of what the intended interpretation was. This is not to say that function words are superfluous. Without them sentences are usually ungrammatical. A sentence like **Nelly went town* which lacks the preposition *to* is not permitted. We have to say *Nelly went to town*.

In English, content words have this property: one of their syllables is more prominent than the rest because it receives MAIN STRESS. This is seen in the words below where the syllable with main stress is preceded by ‘ˈ’:

[2.5]

Initial stress

Medial stress

Final stress

in different positions in the same word. The spelling *kn* is pronounced /kn/ at the end of a word, as in /belkn/, but at the beginning of a word as in *knee* and *knock* the /k/ is dropped and only the *n* is sounded. Similarly, other stop-plus-nasal combinations like *tn* /tn/ and *dn* /dn/ are allowed at the end of a word (e.g. *bottom* /btm/ and *burden* /b :dn/) but these consonant clusters are not permitted at the beginning of a word. Putative words like */tmɪs/ (**miss*) and */dneɪ/ (**dnell*) are just impermissible. In the spoken language we recognise as English words only those forms that have the right combination of sounds for the position in the word where they occur.

Moreover, even when a sound or combination of sounds is allowed, often a somewhat different pronunciation is used depending on the position in which it occurs in a word. This can be seen in the pronunciation of the /l/ sound in standard British English (RP) in different positions in a word. Compare the initial /l/ with the final /l/ in the following:

[2.7]

<i>Word-initial clear</i>	<i>Word-final dark</i>	<i>Pre-consonantal dark</i>
/l/	l/	l/
labour lead loft	spill smell fulfil	milk salt belt quilt
lend let lick leaf	cool bull sprawl	spoilt colt wild

The /l/ sound is always made with the blade of the tongue against the teeth-ridge, with the sides lowered to allow air to escape. But there is a subtle difference. When /l/ is in word-final position or when it is followed by another consonant (as it is in the last two columns), besides the articulatory gestures mentioned above, the back of the tongue is also simultaneously raised towards the soft palate (or velum). This type of /l/ is called dark or velarised /l/. But when /l/ is at the beginning of a word, no velarisation takes place. This latter type of /l/ is called clear or non-velarised /l/. Thus, the kind of /l/ we hear gives an indication of where in a word it appears.

Do not fail to note the use of square brackets. They are used to enclose ALLOPHONES, i.e. variants of a phoneme. Allophones are different sounds, e.g. [l] and [l̥], that occur in different contexts which all represent the same phoneme /l/.

With regard to spelling too, the situation is not chaotic, although admittedly the relationship between letters and phonemes is not always straightforward, as *knee* being pronounced /ni:/ demonstrates. We recognise as English words only those orthographic words that conform to the spelling conventions of English. If, for example, you saw the word *zvroglen* you would treat it as a foreign word. The letter combination *zvr* is not English. There is no way a word in English could start with those letters.

Let me summarise. One sense in which we use the term 'word' is to refer to WORD-FORMS. If we are thinking of the written language, our word-forms are ORTHOGRAPHIC words. These are easily recognised. They normally have a space before and after them. By contrast, in normal spoken language our word-forms are PHONOLOGICAL words. These are more difficult to identify because they are not discrete entities that can be neatly picked off one by one. None the less, phonological words can be identified on the basis of their phonological characteristics such as stress and phonotactic properties.

2.2.2

Words as vocabulary items

We need to distinguish between words in the sense of word-form as opposed to words as vocabulary items. Let us revisit the examples in [2.2.1] on pp. 11–12. If we are considering wordforms, we can see that the

hyphenated word-form *street-lamp* occurs three times. So if we were counting different word-forms, we would count *street-lamp* three times. However, if we were counting distinct words, in the sense of distinct VOCABULARY ITEMS we would only count it once.

The distinction between word-forms and vocabulary items is important. Very often, when we talk about words what we have in mind is not word-forms, but something more abstract—what we will refer to here as LEXEMES (i.e. vocabulary items). Anyone compiling a dictionary lists words in this sense. So, although the word-forms in each of the columns in [2.8] below are different, we do not find each one of them given a separate entry in an English dictionary. The first word in each column is listed under a heading of its own. The rest may be mentioned under that heading, if they do not follow a regular pattern of the language—e.g. *write, written* (past participle), *wrote* (past tense). But if they do follow the general pattern (e.g. *washes, washing, washed; smile, smiling, smiled*) they will be left out of the dictionary altogether. Instead, the grammar will be expected to provide a general statement to the effect that verbs take an *-ing* suffix, which marks progressive aspect, and an *-ed* suffix that marks both the past tense and the past participle, and so on.

[2.8]

WASH	TAKE	BRING	WRITE
wash	take	bring	write
washes	takes	brings	writes
washing	taking	bringing	writing
washed	took	brought	wrote
washed	taken	brought	written

In [2.8] each lexeme (i.e. vocabulary item) that would be entered in a dictionary is shown in capital letters and all the different word-forms belonging to it are shown in lower-case letters.

The examples in [2.8] are all verbs. But, of course, lexemes can be nouns, adjectives or adverbs as well. In [2.9] you will find examples from these other word classes.

[2.9]

	<i>Noun</i>	<i>Adjective</i>	<i>Adverb</i>
a.	MATCH	KIND	SOON
	match	kind	soon
	matches	kinder	sooner
b.	GOOSE	BAD	WELL
	goose	bad	well
	geese	worse	better

In [2.9] we have three pairs of lexemes: the nouns, *match* and *goose*; the adjectives *kind* and *bad*; and adverbs *soon* and *well*. In each case the word-forms belonging to each lexeme in [2.9a] follow a general pattern for words of their type and need not be listed in the dictionary. But all the ones in [2.9b] are irregular and must be listed in the dictionary.

The lexeme is an abstract entity that is found in the dictionary and that has a certain meaning. Word-forms are the concrete objects that we put down on paper (orthographic words) or utter (phonological words) when we use language. The relationship between a lexeme and the word-forms belonging to it is one

It could be a *bat* with which you play cricket or a small, flying mammal. This is a case of LEXICAL AMBIGUITY. We have in this sentence a word-form that represents more than one lexeme with a meaning that is quite plausible. It is not possible to determine the right interpretation of the sentence without looking at the wider context in which it appears.

We have established that the relationship between a word-form and the meaning that it represents is a complex one. This is exploited not only in literature and word-play as we saw above but also in the language of advertising. For instance, a recent British Gas newspaper advertisement for gas heating said:

[2.15]

You will warm to our credit. It's free.

This advertisement exploits the lexical ambiguity that is due to the fact that *warm (to)* can mean 'become enthusiastic' or 'experience a rise in temperature'. Next time you look at an advertisement, see whether it exploits any of the relationships between lexemes and word-forms that we have examined.

2.2.3

Grammatical words

Finally, let us consider the word from a grammatical perspective. Words play a key role in syntax. So, some of their properties are assigned taking into account syntactic factors. Often words are required to have certain properties if they serve certain syntactic purposes. Thus, although in [2.16a] we have the same sense of the same lexeme (*play*) realised by the same word-form (*played*), we know that this word does at least two quite different grammatical jobs in the sentence of which it is a part:

[2.16]

a. She played the flute.

b. She took the flute.

She has played the flute.

She has taken the flute.

If you compare the sentences in [2.16] above, you will see that in [2.16a] the verb *play* is realised by the word-form *played* regardless of whether it simply indicates that the action happened in the past as in the first example or that an action was (recently) completed as in the second example. Contrast this with the situation in [2.16b] where these two grammatical meanings are signalled by two different forms. *Took* indicates that the action happened in the past while *taken* (after *has/had*) indicates that the action is complete. In *She played the flute* and *She took the flute* the words *played* and *took* are described grammatically as the 'past tense forms of the verbs *play* and *take*'. By contrast, in *She has played the flute* and *She has taken the flute* we describe *played* and *taken* as the 'past participle' of *play* and *take*.

Linguists use the term SYNCRETISM to describe situations such as that exemplified by *played* where the same word-form of a lexeme is used to realise two (or more) distinct grammatical words that are represented separately in the grammatical representations of words belonging to some other comparable lexemes. The phenomenon of syncretism is one good reason for distinguishing between word-forms and grammatical words. It enables us to show that words belonging to the same lexeme and having the same form in speech and writing can still differ.

A further example should make the ideas of grammatical words and syncretism even clearer. Consider the verbs in the following sentences:

14 WHAT IS A WORD?

- c. *fair*: *fair* (Adjective) 'beautiful, attractive'
fair (Noun) 'holiday'
-

By contrast, word-forms may have the same pronunciation but different spellings and meanings. Such forms are called HOMOPHONES. See this example from a joke book:

[2.12]

Why does the pony cough?
Because he's a little hoarse.
(Young and Young 1981:57)

The joke is a pun on /h:z/, the pronunciation of the two lexemes represented in writing by *horse* and *hoarse*. Other examples of homophones include *tail* ~ *tale*, *sail* ~ *sale*, *weather* ~ *whether*, *see* ~ *sea*, *read* ~ *reed*, *reel* ~ *real*, *seen* ~ *scene*, *need* ~ *knead*.

Conversely, it is also possible to have several closely related meanings that are realised by the same word-form. The name for this is POLYSEMY. Often you find several senses listed under a single heading in a dictionary. For instance, under the entry for the noun *force*, the *OED* lists over ten senses. I have reproduced the first six below:

[2.13]

1. Physical strength. Rarely in *pl.* (= Fr. *forces*—1818.)
 2. Strength, impetus, violence, or intensity of effect ME.
 3. Power or might; esp. military power ME. b. In early use, the strength (of a defensive work etc.). Subseq., the fighting strength of a ship. 1577.
 4. A body of armed men, an army. In *pl.* the troops or soldiers composing the fighting strength of a kingdom or a commander ME. b. A body of police; often absol. *the force*=policemen collectively. 1851.
 5. Physical strength or power exerted on an object; esp. violence or physical coercion. ME.
 6. Mental or moral strength. Now only, power of effective action, or of overcoming resistance. ME.
-

The line that separates polysemy from homonymy is somewhat blurred because it is not altogether clear how far meanings need to diverge before we should treat words representing them as belonging to distinct lexemes. In [2.13], it is not entirely clear that the sixth sense of the noun *force* is not sufficiently removed from the other meanings to merit an entry of its own. The other meanings all show a reasonably strong family resemblance. But mental or moral strength shows a somewhat weaker relationship.

In the *OED*, there is a separate entry for the lexeme *force*, the verb. It is considered a different lexeme because it has a different meaning and belongs to a different word-class, being a verb and not a noun. Belonging to different word-classes is an important consideration in determining whether separate dictionary entries are needed.

In real-life communication, the lack of a one-to-one match between lexemes and word-forms does not necessarily cause ambiguity. In context, the relevant meaning is normally easy to determine. But there are cases where it is not. For instance, the homonymy of *bat* in [2.14] can cause semantic confusion:

[2.14]

I saw a *bat* under the tree.

[2.10]

of REALISATION or REPRESENTATION or MANIFESTATION. If we take the lexeme *write* which is entered in the dictionary, for example, we can see that it may be realised by any one of the word-forms *write*, *writes*, *writing*, *wrote* and *written* which belong to it. These are the actual forms that are used in speech or appear on paper. When you see the orthographic words *written* and *wrote* on the page, you know that although they are spelt differently they are manifestations of the same vocabulary item WRITE.

The distinction between word-forms and lexemes which I have just made is not abstruse. It is a distinction that we are intuitively aware of from an early age. It is the distinction on which word-play in puns and in intentional ambiguity in everyday life depends. At a certain period in our childhood we were fascinated by words. We loved jokes—even awful ones like [2.10]

The humour, of course, lies in recognising that the word-form *shrimp* can belong to two separate lexemes whose very different and unrelated meanings are none the less pertinent here. It can mean either ‘an edible, long, slender crustacean’ or ‘a tiny person’ (in colloquial English). Also, the word *serve* has two possible interpretations. It can mean ‘to wait upon a person at table’ or ‘to dish up food’. Thus, word-play exploits the lexical ambiguity arising from the fact that the same word-form represents two distinct lexemes with very distinct meanings.

In real-life communication, where potential ambiguity occurs we generally manage to come to just one interpretation without too much difficulty by selecting the most appropriate and RELEVANT interpretation in the situation. Suppose a 20-stone super heavyweight boxer went to Joe’s Vegetarian Restaurant and asked the waiter for a nice shrimp curry and the waiter said in reply, ‘We don’t serve shrimps’, it would be obvious that it was shrimps in the sense of crustaceans that was intended. If, on the other hand, a little man, barely 5 feet tall and weighing a mere 7 stone, went to a fish restaurant and saw almost everyone at the tables around him tucking into a plateful of succulent shrimps, and thought that he would quite fancy some himself, he would be rightly offended if the waiter said ‘We do not serve shrimps.’ It is obvious in this situation that shrimps are on the menu and are dished up for consumption. What is not done is serve up food to people deemed to be puny.

Puns are not restricted to jokes. Many advertisements like that for Standens rely on puns for their effect. Given the context, it is obvious that *sound* is meant to be read in more than one sense here.

Serious literature also uses this device. For instance, the First World War poet Siegfried Sassoon gives the title ‘Base details’ to the poem in which he parodies cowardly generals who stay away at the base, at a safe distance from the action, and gladly speed young soldiers to their death at the front. The word-form *base* in the title represents two distinct lexemes here whose meanings are both relevant: (i) *Base details* are details of what is happening at the *base* (Noun) (meaning ‘military encampment’), and (ii) *Base details* are particulars of something that is *base* (Adjective) (meaning ‘reprehensibly cowardly, mean etc.’).

The term HOMONYM is used to denote word-forms belonging to distinct lexemes that are written and pronounced in the same way. There are separate dictionary entries for such words. *Shrimp* and *base* are examples of homonyms. But perhaps they are not so obvious. Better examples of homonyms are shown in [2.11].

[2.11]

- a. *bat*: *bat* (Noun) ‘a small flying mammal’
 bat (Noun) ‘a wooden implement for hitting a ball in cricket’
- b. *bar*: *bar* (Noun) ‘the profession of barrister’
 bar (Noun) ‘a vertical line across a stave used to mark metrical accent in music’
 bar (Verb) ‘to obstruct’

16 WHAT IS A WORD?

[2.17]

- | | |
|----------------|---|
| a. You hit me. | (=you hit me some time in the past)
or
(=you hit me habitually) |
| b. You cut it. | (=you cut it some time in the past)
or
(=you cut it habitually) |
-

As the paraphrases show, the word-form *hit* belonging to the lexeme *hit* can represent either the present tense or the past tense form of the verb. In other words, there is syncretism. We have two different grammatical words *hit* [+verb, +present] and *hit* [+verb, +past] but a single word-form. The same analysis also applies to *cut*. It can represent either the present or past tense of the verb *cut*.

Syncretism is not limited to verbs. It can apply to other word classes (e.g. nouns) as well:

[2.18]

- (a) The wolf killed a sheep and one deer.
(b) The wolf killed two sheep and three deer.
-

In these two sentences, although the word-form *sheep* belongs to the same lexeme and is unchanged in form, we know that its grammatical value is not the same. In [2.18a] it realises the word with the grammatical properties of noun and singular, but in [2.18b] it represents a plural form. Likewise, the same word-form *deer* represents a singular noun in [2.18a] and a plural noun in [2.18b].

What can we say about the word as an entity that functions as a grammatical unit in the syntax of a language? As mentioned already, the (grammatical) word is normally defined as the MINIMAL FREE FORM that is used in the grammar of a language. Let us now put some flesh on this terse and somewhat cryptic statement.

By free form we mean an entity that can stand on its own and act as a free agent; it is an element whose position in a sentence is not totally dictated by other items. In order to explain what 'freedom' means in this context, we need to take on board two ancillary ideas: POSITIONAL MOBILITY and STABILITY. Although words are not the smallest grammatical units used to construct sentences (see the discussion of morphemes in the next chapter), at the level of sentence organisation the rules of sentence formation treat words as unanalysable units. Often it is possible to change the order in which words appear in a sentence and still produce a well-formed sentence. Words enjoy considerable positional mobility. However, the elements inside a word do not enjoy such mobility. While syntactic rules can transport words to new places in a sentence, they cannot shift in the same way elements that are found inside words. Moving words around in the following produces grammatical sentences with basically the same meaning, but with somewhat different emphasis:

[2.19]

- a. This old industrialist revisited Lancaster, fortunately,
b. Fortunately, this old industrialist revisited Lancaster,
c. Lancaster, this old industrialist revisited, fortunately,
d. Fortunately, Lancaster was revisited by this old industrialist.
-

Evidently, the position of words in a sentence is not rigidly fixed. They can, and often do, get moved around if the communicative needs of the speaker or writer require it. However, the interior of a word is a no-go area for syntactic rules. They are strictly barred from manipulating elements found inside a word. As far as syntax is concerned, words are indivisible units that cannot be split and whose internal units are inaccessible (cf. Bauer 1988, Matthews 1991, Lyons 1968, Di Sciullo and Williams 1987).

The word as a grammatical unit shows stability (or INTERNAL COHESION). The order of elements inside a word is rigidly fixed. If the elements of a sentence are shifted, certain meaningful units (in this case *re-visit-ed* and *fortun-ate-ly*) all move *en bloc*, and their order always remains unchanged. The internal structure of the word cannot be tampered with. We are not allowed to perform operations that would yield words like **ed-visit-re*, **ate-fortune-ly* etc. We will return to this point on p. 33 below.

The definition of the word includes the term 'minimal' for a good reason. This is intended to separate words from phrases like *this old industrialist*. Like words, phrases can occur in isolation and they can be moved from one position to another (as we have seen in [2.19]). But the expression *this old industrialist* is not a minimal form since it contains smaller forms capable of occurring independently namely, *this*, *old* and *industrialist*. Furthermore, the sequence *this old industrialist* does not have the kind of internal cohesion found in words. It can be interrupted by other words e.g. *this wealthy old industrialist*; *this very wealthy, old, benevolent industrialist*.

The assumption that the grammatical word is 'a minimum free form' works well as a rule of thumb. But it encounters difficulties when confronted by a COMPOUND WORD like *wheelbarrow* which contains the words *wheel* and *barrow* which can stand alone. In such cases it is clear that the word is not the smallest meaningful unit that can be used on its own. It is for this reason that the definition of the word as the unit on which purely syntactic operations can be performed is preferable. In the case of compounds this definition works. The interior of a compound is a syntactic no-go area. Syntactic rules are not allowed to apply separately to words that make up a compound. Thus, for example although the nouns *wheel* and *barrow* can be modified by the adjective *big* (*[big barrow]*, *[big wheel]*), and although we can talk of *[big wheelbarrow]*, in which case *big* modifies the entire compound, there is no possibility of saying *wheel [big barrow]*, with the adjective only modifying the second element of the compound word.

2.3

SUMMARY

In this chapter we have established that normally, the term 'word' is used ambiguously. To avoid the ambiguity, we need to distinguish between three different types of word: (i) a word-form (i.e. a particular physical manifestation of one or more lexemes in speech or writing); (ii) a vocabulary item (i.e. lexeme); and (iii) a unit of grammatical structure that has certain morphological and syntactic properties.

We will revisit the distinction between lexemes, grammatical words and word-forms mainly in Chapters 7 and 11. In Chapter 7 our main concern will be the realisation of words in speech and in writing. In Chapter 11 we will show that this distinction is not an artefact of the linguist's analysis. Rather, it is a distinction that is well supported by studies in the way in which we store words in the mind and retrieve them for use in communication in real life.

In the coming chapters, in cases where the relevant sense of the term 'word' is clear from the context I will not spell out whether it is the word as a vocabulary item, grammatical word, phonological or orthographic form that is being dealt with. But where it is not clear, I will indicate the sense in which I am using this term. We are now in a position to consider in detail the internal structure of words. That is the task of the next chapter.

[3.2]

childish hopeless sooner mended elephants re-boil unsafe ex-wife

You would have to give a different answer. You would need to tell your interrogator, who by now would be getting increasingly bewildered, that the words in [3.2] can be divided into smaller units of meaning as shown in [3.3]:

[3.3]

child-*ish* hope-*less* soon-*er* mend-*ed* elephant-*s* re-boil *un*-safe ex-wife

The part of the word that is not italicised can function as an independent word in the grammar. Indeed, each of the nonitalicised chunks is a word (i.e. vocabulary item) that is listed as such in the dictionary. By contrast, the italicised bits, though meaningful (and their meanings can be indicated as shown in [3.4]), cannot function on their own in the grammar.

[3.4]

-ish	'having the (objectionable) qualities of'	child- <i>ish</i> = 'having the qualities of a child'
-less	'without X'	hopeless= 'without hope'
-er	'more X'	sooner= 'more soon'
-ed	'past'	mended= 'mend in the past'
-s	'plural'	elephants= 'more than one elephant'
re	'again'	re-boil= 'boil again'
un	'not X'	unsafe= 'not safe'
ex	'former'	ex-wife= 'former wife'

What we have done to the words in [3.4] can be done to thousands of other words in English. They can be decomposed into smaller units of meaning (e.g. *re-* 'again') or grammatical function (e.g. *-ed* 'past').

The term MORPHEME is used to refer to the smallest unit that has meaning or serves a grammatical function in a language. Morphemes are the atoms with which words are built. It is not possible to find sub-morphemic units that are themselves meaningful or have a grammatical function. Thus, given *-less* or *un-*, it would make no sense to try to assign some identifiable meaning to any part of these forms. Of course, it is possible to isolate the individual sounds /l-ɪ-s/ or /-n/, but those sounds in themselves do not mean anything.

We have now established that words are made up of morphemes. But how do we recognise a morpheme when we see one? Our definition of the morpheme as the smallest unit of meaning (or grammatical function) will be the guiding principle. Any chunk of a word with a particular meaning will be said to represent a morpheme. That is how we proceeded in [3.3] and [3.4] above.

Morphemes tend to have a fairly stable meaning which they bring to any word in which they appear. If we take *re-* and *un-*, for example, they mean 'again' and 'not' respectively—not just in the words we have listed above, but also in thousands of other words. Usually morphemes are used again and again to form different words. Thus *re-* meaning 're-do whatever the verb means' can be attached before most verbs to yield a new word with a predictable meaning (e.g. *re-run*, *re-take*, *re-build* etc.). In like manner, *un-* meaning 'not X' (where X stands for whatever the adjective means) can be attached to various adjectives (e.g. *un-real*, *un-clean*, *un-happy* etc.) to yield a new word with a predictable negative meaning.

Chapter 3

Close encounters of a morphemic kind

3.1

THE QUEST FOR VERBAL ATOMS

We saw in the last chapter that the word is the smallest meaningful unit of language that can function independently in the grammar. A word can be used on its own, without appending it to some other unit. Thus, in the word *childish* we can isolate *child* and use it on its own because it is a word in its own right. But we cannot use *-ish* as a stand-alone unit, for *-ish* is not a word.

While recognising that words are the smallest meaningful units which function independently in the grammar, we also need to recognise that words can be decomposed into smaller units that are also meaningful. Our task in this chapter is to explore the internal structure of words in order to gain some understanding of the basic units which are used to form words.

3.2

CLOSE MORPHOLOGICAL ENCOUNTERS: ZOOMING IN ON MORPHEMES

Originally 'morphology' meant the study of biological forms. But nineteenth-century students of language borrowed the term and applied it to the study of word-structure. In linguistics MORPHOLOGY is the study of the formation and internal organisation of words.

Let us begin our morphological analysis by considering half a dozen words (not altogether randomly chosen):

[3.1]

hope soon mend boil safe leaf word elephant

Obviously all the words in [3.1] have a meaning, but lack internal structure. We cannot identify any smaller units that are themselves meaningful which occur inside them. If a Martian stopped you in a street near the local zoo and enquired what *phant* in *elephant* or *ho* in *hope* means, you would think she was asking a most bizarre question that did not merit an answer. Or you might condescendingly explain that, of course, in each case the whole word means something, but its parts cannot be said to mean anything on their own. Though somewhat puzzled, the Martian might accept your explanation.

But, being the persistent type, let us suppose she enquired further whether the words in [3.2] were also indivisible into smaller meaningful units:

The segmentation of words into morphemes is not a trivial and arcane pastime indulged in by linguists to while away the time on a wet Bank Holiday afternoon. It is something that is important for all users of language. During your lifetime, you will probably encounter hundreds of thousands of different words. Many of these words will be new to you. For no matter how extensive your vocabulary is, you will inevitably come across words that are unfamiliar. It is impossible for anyone to know all the words that are found in English.

So, what do you do when faced with an unfamiliar word? Reach for a good dictionary? Perhaps. But this is not always feasible. Nor is it always necessary. Very often you just figure out what the strange word means using the context, together with your knowledge of the meaning of the morphemes which the word contains. You normally do this subconsciously. What we are doing here is making explicit your tacit knowledge of word-structure.

Imagine this scenario. In 1992, a newspaper report on the war in the Bosnian republic states that what we are witnessing is the *Lebanonisation* of Bosnia. Suppose you have not encountered the word *Lebanonisation* before. Would you understand what the writer is saying? Probably you would—without looking it up in any dictionary. How would you do it? The answer is simple. By using your knowledge of the world—in particular history (*Balkanisation*)—and your knowledge of current affairs (the civil war in Lebanon) plus your knowledge of the principles of word-formation you are able to work out the meaning of *Lebanonisation*.

Let us focus on principles of word-formation. You know that *-ize/-ise* is used when talking about nations to mean 'to make X', e.g. from *America* we get *Americanise*, from *Korea* we get *Koreanise*, from *Kenya* we get *Kenyanise* etc. By attaching *-(an)ise* we turn a noun into a verb. So, given the noun *Lebanon* we can form the verb *Lebanonise*. Next, from the verb *Lebanonise*, we can create a new noun by adding *-ation* (which forms nouns of action).

If you know that various warlords created warring fiefdoms that destroyed the Lebanese state during the civil war that raged in Lebanon in the 1970s and 1980s, you will know that the Croats, Muslims and Serbs engaged in the Bosnian conflict risk doing the same to the Bosnian state in the 1990s. *Lebanonisation* is the act of 'turning a country into another Lebanon'. Thus, our knowledge of word-structure contributes to our understanding of the meaning of unfamiliar words.

We have demonstrated that words can be decomposed into morphemes. Now we are going to see that words have INTERNAL STRUCTURE. A simple way of showing this is to analyse words like *uncanny* and *unhappy*. From these words we can derive *uncannier* and *unhappier*. If you analyse *unhappier*, you will see that extracting the correct meaning 'more [not happy]' (i.e. sadder) rather than the incorrect one 'not [more happy]' (i.e. not happier) depends on the way we group together the morphemes. In the first analysis where *unhappier* is interpreted as *sadder*, the meaning 'not' conveyed by *un-* is bracketed together with *happy* [unhappy] as one unit and this is intensified by the *-er* suffix. In the alternative second analysis, *happy* and *-er* are bracketed together as a unit [happier] (i.e. more happy) which then is negated by [un-] to give 'not more happy', which is incorrect. When someone is *unhappier*, it does not mean they are simply less happy, it means rather that they are not happy at all. They are sad. This shows that morphemes in a word with several morphemes may be grouped together in different ways for semantic purposes. The way in which this is done has semantic consequences. Conceivably, morphemes could be thrown together higgledy-piggledy to form a word. So long as you had the right morphemes, a well-formed word would pop out. But that is definitely not the case. Words have internal structural groupings, as we have seen.

Furthermore, the sequencing of morphemes in a word may be subject to restrictions. Take a word like *ungovernability* which contains four morphemes, namely *un-*, *govern*, *abil*, *ity*. Everyone who knows this

[2.17]

- | | |
|----------------|---|
| a. You hit me. | (=you hit me some time in the past)
or
(=you hit me habitually) |
| b. You cut it. | (=you cut it some time in the past)
or
(=you cut it habitually) |
-

As the paraphrases show, the word-form *hit* belonging to the lexeme *hit* can represent either the present tense or the past tense form of the verb. In other words, there is syncretism. We have two different grammatical words *hit* [+verb, +present] and *hit* [+verb, +past] but a single word-form. The same analysis also applies to *cut*. It can represent either the present or past tense of the verb *cut*.

Syncretism is not limited to verbs. It can apply to other word classes (e.g. nouns) as well:

[2.18]

- (a) The wolf killed a sheep and one deer.
(b) The wolf killed two sheep and three deer.
-

In these two sentences, although the word-form *sheep* belongs to the same lexeme and is unchanged in form, we know that its grammatical value is not the same. In [2.18a] it realises the word with the grammatical properties of noun and singular, but in [2.18b] it represents a plural form. Likewise, the same word-form *deer* represents a singular noun in [2.18a] and a plural noun in [2.18b].

What can we say about the word as an entity that functions as a grammatical unit in the syntax of a language? As mentioned already, the (grammatical) word is normally defined as the MINIMAL FREE FORM that is used in the grammar of a language. Let us now put some flesh on this terse and somewhat cryptic statement.

By free form we mean an entity that can stand on its own and act as a free agent; it is an element whose position in a sentence is not totally dictated by other items. In order to explain what 'freedom' means in this context, we need to take on board two ancillary ideas: POSITIONAL MOBILITY and STABILITY. Although words are not the smallest grammatical units used to construct sentences (see the discussion of morphemes in the next chapter), at the level of sentence organisation the rules of sentence formation treat words as unanalysable units. Often it is possible to change the order in which words appear in a sentence and still produce a well-formed sentence. Words enjoy considerable positional mobility. However, the elements inside a word do not enjoy such mobility. While syntactic rules can transport words to new places in a sentence, they cannot shift in the same way elements that are found inside words. Moving words around in the following produces grammatical sentences with basically the same meaning, but with somewhat different emphasis:

[2.19]

- a. This old industrialist revisited Lancaster, fortunately,
b. Fortunately, this old industrialist revisited Lancaster,
c. Lancaster, this old industrialist revisited, fortunately,
d. Fortunately, Lancaster was revisited by this old industrialist.
-

word knows that these four morphemes must appear in the order in [3.5a]. Any other order is strictly forbidden:

[3.5]

- a. un-govern-abil-ity
 - b. *govern-abil-un-ity
 - c. *ity-un-abil-govern
 - d. *abil-un-ity-govern
 - e. *un-govern-ity-abil etc.
-

Clearly, knowing a word means not just knowing the morphemes it contains, but also the rigid order in which they are allowed to appear. We will return to this point in section (4.4).

To sum up the discussion so far, words are built using morphemes. If we know how morphemes are used to form words, we do not need to be unduly flustered when we come across a strange word. Usually it is possible to work out the meaning of a strange word if it contains familiar morphemes.

3.3

MORPHEMES AND THEIR DISGUISES

The identification of morphemes is not altogether straightforward. This is because there is no simple one-to-one correspondence between morphemes and the speech sounds that represent them. In this section we will attempt to unravel the complexities of the relationship between morphemes and the actual forms (sounds of groups of sounds) by which they are manifested in speech.

3.3.1

Allomorphs: morph families

Any physical form that represents a morpheme is called a MORPH. The forms *-ish*, *-less*, *-er*, *-ed*, *-s*, *re-*, *un-* and *ex-* in [3.4] on p. 31 are all morphs. Morphological analysis begins with the identification of morphs, i.e. forms that carry some meaning or are associated with some grammatical function. In *asparagus* there is just one morph but in all the words in [3.4] there are two.

It is important not to confuse morphs with SYLLABLES. When we talk of morphs we have in mind sounds that can be related to a particular meaning or grammatical function (e.g. plural or past tense). However, when we talk of syllables all we have in mind are chunks into which words can be divided for the purposes of pronunciation.

This is not an abstruse distinction. We are not being pedantic. It is a distinction that matters to ordinary people because human languages are organised in such a way that the construction of units that are meaningful is normally in principle separate from the construction of strings that are pronounceable. Thus, for rhythmical effect, nursery rhymes often use nonsense syllables like '*Deedle, deedle*' in '*Deedle deedle dumpling my son John*' which do not represent anything meaningful.

Alternatively, a sound representing a morpheme may not be a syllable in its own right, e.g. by itself, the *-s* which represents the plural morpheme is not a syllable. The word *cats* has two morphemes, *cat* and *-s*, but it is all just one syllable. The single syllable *cats* realises two morphemes. The converse situation, where

several syllables realise a single morpheme, is equally possible. Thus, the trisyllabic and quadrisyllabic word-forms *elephant* and *asparagus* both realise just a single morpheme.

The nature of the relationship between sounds and morphemes is intriguing. At first sight, it might look reasonable to assume that morphemes are made up of PHONEMES. We might be tempted to think that *cat*, the English morpheme with the meaning is made up of the phonemes /kæt/. But we have several kinds of evidence showing that this is not the case.

First, if morphemes were *made up* of phonemes, a given morpheme would be uniquely associated with a given phonological representation. In reality, the same morpheme can be realised by different morphs (i.e. sounds or written forms). Morphs which realise the same morpheme are referred to as ALLOMORPHS of that morpheme.

The INDEFINITE ARTICLE is a good example of a morpheme with more than one allomorph. It is realised by the two forms *a* and *an*. The sound at the beginning of the following word determines the allomorph that is selected. If the word following the indefinite article begins with a consonant, the allomorph *a* is selected, but if it begins with a vowel the allomorph *an* is used instead:

[3.6]

a.	a dictionary	b.	an island
	a boat		an evening
	a pineapple		an opinion
	a leg		an eye
	a big (mat)		an old (mat)
	a dull (song)		an exciting (finish)

Hence the incorrectness of the sentence marked with an asterisk in [3.7]:

[3.7]

- | | |
|----|---------------------------------------|
| a. | I spent <i>an</i> evening with them. |
| | *I spent <i>a</i> evening with them. |
| b. | I spent <i>the</i> evening with them. |
-

Allomorphs of the same morpheme are said to be in COMPLEMENTARY DISTRIBUTION. This means that they do not occur in identical contexts and therefore they cannot be used to distinguish meanings. In other words, it is impossible to have two otherwise identical utterances that differ in their meanings depending on the allomorph of a morpheme that is selected. So, because *a* and *an* both realise the same indefinite article morpheme, it is impossible to have two sentences like those in [3.7a] above which are identical in all ways, except in the choice of *a* or *an*, but mean different things.

Complementary distribution presupposes the more basic notion of DISTRIBUTION. Distribution is to do with establishing facts about the occurrence of allomorphs of a particular morpheme. It is concerned with establishing the contexts in which the morpheme which we are investigating occurs and the allomorphs by which it is realised in those different contexts. In other words, by distribution we mean the total set of distinct linguistic contexts in which a given form appears, perhaps in different guises. For instance, the indefinite article has the distribution: *a* before consonants (e.g. *a tree*) and *an* before vowels (e.g. *an eagle*).

As mentioned already, such functionally related forms which all represent the same morpheme in different environments are called allomorphs of that morpheme. Another way of putting it is

that allomorphs are forms that are phonologically distinguishable which, none the less, are not functionally distinct. In other words, although they are physically distinct morphs with different pronunciations, allomorphs do share the same function in the language.

An analogy might help to clarify this point. Let us compare allomorphs to workers who share the same job. Imagine a jobshare situation where Mrs Jones teaches maths to form 2DY on Monday afternoons, Mr Kato on Thursday mornings and Ms Smith on Tuesdays and Fridays. Obviously, these teachers are different individuals. But they all share the role of 'maths teacher' for the class and each teacher only performs that role on particular days. Likewise, all allomorphs share the same function but one allomorph cannot occupy a position that is already occupied by another allomorph of the same morpheme. To summarise, we say that allomorphs of a morpheme are in complementary distribution. This means that they cannot substitute for each other. Hence, we cannot replace one allomorph of a morpheme by another allomorph of that morpheme and change meaning.

For our next example of allomorphs we will turn to the plural morpheme. The idea of 'more than one' is expressed by the plural morpheme using a variety of allomorphs including the following:

[3.8]

	<i>Singular</i>	<i>Plural</i>
a.	rad-ius cactus	radi-i cact-i
b.	dat-um strat-um	dat-a strat-a
c.	analys-is ax-is	analys-es ax-es
d.	skirt road branch	skirt-s road-s branch-es

Going by the orthography, we can identify the allomorphs *-i*, *-a*, *-es* and *-s*. The last is by far the commonest: see section (7.3).

Try and say the batch of words in [3.8d] aloud. You will observe that the pronunciation of the plural allomorph in these words is variable. It is [s] in *skirts*, [z] in *roads* and [ɪz] (or for some speakers [eɪz]) in *branches*. What is interesting about these words is that the selection of the allomorph that represents the plural is determined by the last sound in the noun to which the plural morpheme is appended. We will return to this in more depth in section (5.2).

We have already seen, that because allomorphs cannot substitute for each other, we never have two sentences with different meanings which solely differ in that one sentence has allomorph X in a slot where another sentence has allomorph Y. Compare the two sentences in [3.9]:

[3.9]

a.	They have two cats [eɪ hæv tu: kæt-s] *[eɪ hæv tu: kæt-z]	b.	They have two dogs [eɪ hæv tu: dg-z] *[eɪ hæv tu: dg-s]
----	---	----	---

Those morphemes that are allowed to occur on their own in sentences as words are called FREE MORPHEMES while those morphemes that must occur in the company of some other morphemes are called BOUND MORPHEMES. In [3.12] the bound morphemes are italicised.

[3.12]

pest	<i>pest(i)-cide</i>
modern	<i>post-modern-ist</i>
child	<i>child-ish</i>
pack	<i>pre-pack-ed</i>
laugh	<i>laugh-ing</i>

The free morphemes in [3.12] can all be manipulated by syntactic rules; they can stand on their own as words. By contrast, it is impossible to use the forms *-cide*, *post-*, *-ist*, *-ish*, *pre-*, *-ed* or *-ing*, independently.

So far, all the examples of free morphemes that function as roots that we have encountered have been content words (see p. 14). However, not all free morphemes are content words. Some are employed to indicate grammatical functions and logical relationship rather than to convey lexical or cognitive meaning in a sentence. Hence such words are called FUNCTION WORDS. They include words such as the following:

[3.13]

articles:	<i>a/an, the</i>
demonstratives:	<i>e.g. this, that, these and those</i>
pronouns:	<i>e.g. I, you, we, they, my, your, his, hers, who etc.</i>
prepositions:	<i>e.g. in, into, on to, at, on etc.</i>
conjunctions:	<i>e.g. and, or, but, because, if etc.</i>

In ordinary language use such words are extremely common. But on their own they would not convey a lot of information. If you received a telegram like *But it my on to the in* you might suspect that the sender either had a strange sense of humour or was not mentally sound.

3.5

SOUND SYMBOLISM: PHONAESTHEMES AND ONOMATOPOEIA

In the vast majority of words, the relationship between sound and meaning is arbitrary (see p. 2). There is no reason why a particular sound, or group of sounds, should be used to represent a particular word, with a particular meaning. If someone asked you what [b] in *bed* or [str] in *strange* meant, you would think they were asking a very odd question. As a rule, sounds *qua* sounds do not mean anything.

However, the general principle that says that the link between sound and meaning in words is arbitrary is occasionally dented. This happens in two sets of circumstances. First, certain individual sounds, or groups of sounds, which do not represent a specific enough meaning to be called morphs nevertheless appear to be vaguely associated with some kind of meaning. Such sounds are called PHONAESTHEMES.

As our first example of a phonaestheme, let us take the RP vowel [ɒ] (which is historically descended from [U], the vowel that is still used in words like *dull* and *hut* in the north of England). This phonaestheme is found in words associated with various kinds of dullness or indistinctness, e.g. *dull, thud, thunder, dusk,*

table as a container for cut fresh flowers. She calls hers a flagon, for that is what she is using it as. You call yours a vase.

Here are the questions now: are these objects 'flagons' or 'vases'? Which one of you is right? I am not being evasive if I say that both of you are right. For, although the two objects are identical as far as their form, their physical properties, is concerned, they are very different with regard to the functions that they serve in your two households.

There are numerous linguistic parallels. What is physically the same linguistic form can be used to represent distinct morphemes. In order for forms to be regarded as allomorphs belonging to the same morpheme, it is not sufficient for them to have the same form—to be pronounced or written in the same way. They must also have the same grammatical or semantic function. The significance of this point was hinted at in the discussion of *un-* in *unlocked* and *untidy* when we showed that the same morph can represent different morphemes. It should become even more obvious when you consider the form *-er* in the following:

[3.11]

a.	think – thinker	drive – driver
	write – writer	sing – singer
	sweep – sweeper	sell – seller
b.	cook – cooker	strain – strainer
	receive – receiver	compute – computer
	propel – propeller	erase – eraser
c.	London – Londoner	north – northerner
	Iceland – Icelander	east – easterner
	New York – New Yorker	Highlands – Highlander

The same form, *-er*, represents three different meanings and hence has to be assigned to three distinct morphemes. In [3.11a] it forms an agentive noun from a verb, with the meaning 'someone who does X' (i.e. whatever the verb means). In [3.11b] the same *-er* forms an instrumental noun from a verb, with the meaning 'something used to X' (i.e. to do whatever the verb means). Finally, in [3.11c] the same *-er* form is attached to a noun referring to a place to mean 'an inhabitant of'.

Clearly, the same form does serve different functions here. So, it realises different morphemes. This is further evidence that should quickly disabuse us of the assumption that morphemes are made up of morphs. Not only can a single morpheme have several allomorphs (as in the case of the plural morpheme), the same morph (e.g. *-er*) can represent different morphemes. There is no simple one-to-one matching of morphemes with morphs.

3.4

FREEDOM AND BONDAGE

When we classify morphemes in terms of where they are allowed to appear, we find that they fall into two major groupings. Some morphemes are capable of occurring on their own as words, while other morphemes are only allowed to occur in combination with some other morpheme(s) but they cannot be used by themselves as independent words.

We cannot find two otherwise identical sentences which differ in meaning simply because the word *cats* is pronounced as [kæt-s] and *[kæt-z] respectively. Likewise, it is not possible to have two otherwise identical sentences with different meanings where the word *dogs* is pronounced as [dgz] and *[dgs]. In other words, the difference between the allomorphs [s] and [z] of the plural morpheme cannot be used to distinguish meanings.

3.3.2 Contrast

Different morphemes CONTRAST meanings but different allomorphs do not. If a difference in meaning is attributable to the fact that one minimal meaningful unit has been replaced by another, we identify the morphs involved as manifestations of distinct morphemes. So, in [3.7] on p. 36 the indefinite article realised by *a* or *an* is a distinct morpheme from the definite article realised by *the* since a semantic difference is detectable when *a* or *an* is replaced with *the*.

A further example of contrast is given in [3.10]:

[3.10]

- | | | | |
|----|----------------------|----|----------------|
| a. | I unlocked the door. | b. | She is untidy. |
| | I re-locked the door | | |

The two sentences in [3.10a] mean very different things. Since they are identical except for the fact that where one has *un-* the other has *re-*, the difference in meaning between these two sentences is due to the difference in meaning between the morphemes realised by *re-* (meaning 'do again') and *un-* (meaning 'reverse the action').

Now, contrast the *un-* of *unlocked* with the *un-* of *untidy*. In both cases we have the same morph *un-* (which is spelt and pronounced in exactly the same way). But it is obvious that *un-* represents different morphemes in these two word-forms. In *I unlocked the door* the morph *un-* found in *unlocked* realises a reversive morpheme which is attached to verbs—it reverses the action of locking. But in *untidy* it realises a negative morpheme attached to adjectives—*untidy* means 'not tidy'. (If a person is *untidy*, it does not mean that at some earlier point they were tidy and someone has reversed or undone their tidiness.)

If morphemes were made up of phonemes a simple correlation of morphs with morphemes is what we would find. But, in fact, it is quite common for the same phonological form (i.e. morph) to represent more than one morpheme. It is from the context that we can tell which morpheme it represents. This is the second piece of evidence against the assumption that morphemes are composed of phonemes.

The complex relationship between morphemes and the allomorphs that represent them gives us a window through which we can glimpse one of the most fascinating aspects of language: the relationship between FORM and FUNCTION. In linguistics we explore the form of various elements of language structure, e.g. words and sentences, because it is important to know how they are constructed. However, form is not everything. We are also interested in knowing what linguistic elements are used for, what function they serve.

Just consider for a moment this non-linguistic analogy. Imagine a friend returns from a foreign vacation with two beautiful ornamental glass containers with a globular shape and gives one to you as a present and keeps the other for herself. She does not tell you what your present is used for. She uses hers as a vessel for containing wine at the table—she got the idea of buying these containers when she was served wine in a similar container in a fancy restaurant. You do not know this. You look at your present and decide to put it on the

3.6.2

Get the glue (agglutinating languages)

In an ideal agglutinating language most words contain more than one morpheme and the morphemes are realised by morphs arranged in rows like corn on the cob. The morphs can be neatly picked off, one by one. Swahili is a good example of an agglutinating language as you can see:

[3.16]

- | | |
|----|--|
| a. | ni-ta-pik-a
I-future-cook-BVS
'I will cook' |
| b. | a-li-tu-pik-i-a
s/he-past-us-cook-for-BVS
's/he cooked for us' |
| c. | tu-li-wa-lim-ish-a
we-past-them-cultivate-cause-BVS
'we made them cultivate' |

Note: BVS = basic verbal suffix which is normally *-a*

In a Swahili word, it is normally possible to say which morph represents which morpheme. Most morphs only represent one morpheme at a time and do not FUSE with adjacent morphemes, as say, plural marking does in *leaves* where it is partly signalled by the suffix *-es /-Iz/* and partly by the change of the */f/* of *leaf* to */v/*. It is as if the word is constructed by gluing together separable, discrete morphs.

3.6.3

Labyrinthine words (synthetic languages)

In a SYNTHETIC LANGUAGE a word normally contains more than one morpheme. In this respect synthetic languages resemble agglutinating languages. However, whereas in an agglutinating language the morphemes and the morphs that realise them are arranged in a row one after the other, the morphs of a synthetic language are to a considerable extent fused together and cannot be separated neatly one from the other. Furthermore, the morphemes themselves are not arranged in a row. Rather, they are all thrown together in a big pot like pot-pourri. It is impossible to separate the different strands.

Latin is a classic example of an inflecting language. Any attempt to segment Latin words into morphs in such a way that each morph is associated uniquely with a particular morpheme very soon runs into trouble. You can see this for yourself if you attempt to segment the various word-forms of the nouns *m nsa* 'table' and *fl s* 'flower' into their constituent morphs and try to match those morphs with the corresponding morphemes:

[3.17]

a.	<i>Case</i>	<i>Singular</i>	<i>Plural</i>	b.	<i>Singular</i>	<i>Plural</i>
	Nominative	<i>m nsa</i>	<i>m nsae</i>		<i>fl s</i>	<i>fl r s</i>
	Accusative	<i>m nsam</i>	<i>m nsas</i>		<i>fl rem</i>	<i>fl r s</i>
	Genitive	<i>m nsae</i>	<i>m nsarum</i>		<i>fl ris</i>	<i>fl rum</i>



On the Ning Nang Nong

On the Ning Nang Nong
 Where the cows go Bong!
 And the Monkeys all say Boo!
 There's a Nong Nang Ning
 Where the trees go Ping!
 And the tea pots Jibber Jabber Joo.
 On the Nong Ning Nang
 All the mice go Clang!
 And you just can't catch 'em when they do!
 So it's Ning Nang Nong!
 Cows go Bong!
 Nong Nang Ning!
 Trees go Ping!
 Nong Ning Nang!
 The mice go Clang!
 What a noisy place to belong,
 Is the Ning Nang Ning Nang Nong!
 SPIKE MILLIGAN

blunt, mud, slush, sludge, slump etc. Obviously, the vowel [ʊ] *per se* does not mean 'dull'. If it did, *dim* which contains the vowel [ɪ] would not be a virtual synonym for *dull*.

Many words which mean 'to talk indistinctly' contain one or more occurrences of the labial consonant [m], which is made with the lips firmly closed, preventing clear articulation. That way, the very act of pronouncing the word iconically mimics a key aspect of its meaning. You can see this if you watch yourself in a mirror saying words like *mumble, murmur, mutter, muted, grumble* etc. It is probably not an accident that these words also contain the phonaestheme [m]. Similarly, the sound [mp] (spelled *-ump*) as in *clump, dump, bump, lump* and *hump* is often found at the end of words which are associated with heaviness and clumsiness although no one would wish to suggest that *-ump* in itself represents the ideas of heaviness and clumsiness. Interestingly, here again we have the vowel [ʊ] followed by the labial consonants [mp].

Observe also that whereas [ʊ] tends to have associations of heaviness or dullness, the high front vowels [i:] and [ɪ] frequently occur as phonaesthemes in words associated with smallness, as in *wee, teeny-weeny, lean, meagre, mini, thin* and *little*. (The fact that *big* has the opposite meaning just goes to show that phonaesthemes only represent a tendency.)

Second, and more importantly, in addition to phonaesthemes, there are onomatopoeic words in which a direct association is made between the sounds of a word-form and the meaning that it represents. In cases of ONOMATOPOEIA, the sounds (*qua* sounds and not as morphs) symbolise or reflect some aspect of the meaning of the word that they represent. So, if speakers of any language want an onomatopoeic word for the noise a cat makes, they will not choose a noise like *bimboban*—except, perhaps, in the land of the Ning Nang Nong.

The words for sounds made by various animals e.g. *neigh, miaow, moo* etc. are the most obvious examples of onomatopoeia. But there are others such as *roar, crack, clang, bang, splash, swish, whoosh, buzz, hiss, cheep, bleep, gurgle, plop* and *plod*. In the case of onomatopoeic words, the relationship between sound and meaning is to some extent ICONIC. The sounds mimic an aspect of the meaning of the linguistic sign much in the same way that this iconic sign for a restaurant represents, more or less directly, the meaning 'restaurant'. This symbol is still conventional to some degree. To people who eat with chopsticks, it might not be immediately obvious why this sign represents a restaurant (rather than a cutlery shop), but once it is pointed out the link can be seen quite easily.

Onomatopoeic words are iconic in so far as they directly reflect some aspect of the meaning of what they stand for. So, conventionally in English cows go 'moo' and horses go 'neigh' and bees go 'buzz'. That is why Spike Milligan's nonsense poem 'On the Ning Nang Nong' is bizarre.

To be onomatopoeic, the sound must imitate to some degree an aspect of the noise made by the bird or animal. But exactly what is imitated will vary from language to language. An English cock will say *cockadoodledoo*, a Russian cock *kukuriku* and in Uganda it may say *kookolilookoo*. (These differences are not attributable to dialectical variation among the males of the *Gallus domesticus* species.) Onomatopoeic words are not purely and simply formed by mimicking precisely the meanings that they convey. To some extent onomatopoeic words are also moulded by linguistic convention. That is why in different places in the world different onomatopoeic words may be used for the same animal or bird noise.

3.6

VERBAL BLUEPRINTS

Linguistic theory incorporates the hypothesis that there are universal principles of grammar that regulate the amount of variation in linguistic structure across languages. In the last section we saw the marginal role played by sound symbolism in word-formation. This does not obscure the fact that normally languages form