1 Introduction

1.1 What is morphology?

Morphology is the study of the internal structure of words.[1] Somewhat paradoxically, morphology is both the oldest and one of the youngest subdisciplines of grammar. It is the oldest because, as far as we know, the first linguists were primarily morphologists. The earliest extant grammatical texts are well-structured lists of morphological forms of Sumerian words, some of which are shown in (1.1). They are attested on clay tablets from Ancient Mesopotamia and date from around 1600 BCE.

(1.1)

badu	'he goes away'	*inĝen*	'he went'
baduun	'I go away'	*inĝenen*	'I went'
bašidu	'he goes away to him'	*inšiĝen*	'he went to him'
bašiduun	'I go away to him'	*inšiĝenen*	'I went to him'

(Jacobsen 1974: 53–4)

Sumerian was the traditional literary language of Mesopotamia but, by the second millennium BCE, it was no longer spoken as a medium of everyday communication (having been replaced by the Semitic language Akkadian), so it needed to be recorded in grammatical texts. Morphology was also prominent in the writings of the greatest grammarian of Antiquity, the Indian Pāṇini (fifth century BCE), and in the Greek and Roman grammatical tradition. Until the nineteenth century, Western linguists often thought of grammar as consisting primarily of word structure, perhaps because the

[1] The reader should be aware that this sentence, while seemingly straightforward, conceals a controversy – there is no agreed upon definition of 'word'. The relevant issues are addressed in Chapter 9, but here, and through most of the book, we will appeal to a loose, intuitive concept of 'word'.

classical languages Greek and Latin had fairly rich morphological patterns that were difficult for speakers of the modern European languages.

This is also the reason why it was only in the second half of the nineteenth century that the term *morphology* was invented and became current. Earlier there was no need for a special term, because the term *grammar* mostly evoked word structure, i.e. morphology. The terms *phonology* (for sound structure) and *syntax* (for sentence structure) had existed for centuries when the term *morphology* was introduced. Thus, in this sense morphology is a young discipline.

Our initial definition of morphology, as the study of the internal structure of words, needs some qualification, because words have internal structure in two very different senses. On the one hand, they are made up of sequences of sounds (or gestures in sign language), i.e. they have internal phonological structure. Thus, the English word *nuts* consists of the four sounds (or, as we will say, *phonological segments*) [nʌts]. In general, phonological segments such as [n] or [t] cannot be assigned a specific meaning – they have a purely contrastive value (so that, for instance, *nuts* can be distinguished from *cuts, guts, shuts*, from *nets, notes, nights*, and so on).

But often formal variations in the shapes of words correlate systematically with semantic changes. For instance, the words *nuts, nights, necks, backs, taps* (and so on) share not only a phonological segment (the final [s]), but also a semantic component: they all refer to a multiplicity of entities from the same class. And, if the final [s] is lacking (*nut, night, neck, back, tap*), reference is made consistently to only one such entity. By contrast, the words *blitz*, *box*, *lapse* do not refer to a multiplicity of entities, and there are no semantically related words **blit, *bok, *lap*.[2] We will call words like *nuts* '(morphologically) **complex words**'.

In a morphological analysis, we would say that the final [s] of *nuts* expresses plural meaning when it occurs at the end of a noun. But the final [s] in *lapse* does not have any meaning, and *lapse* does not have morphological structure. Thus, morphological structure exists if there are groups of words that show identical partial resemblances in both form and meaning. Morphology can be defined as in Definition 1.

Definition 1:
Morphology is the study of systematic covariation in the form and meaning of words.

It is important that this form–meaning covariation occurs systematically in groups of words. When there are just two words with partial form–meaning resemblances, these may be merely accidental. Thus, one would

[2] The asterisk symbol (*) is used to mark nonexistent or impossible expressions.

not say that the word *hear* is morphologically structured and related to *ear*. Conceivably, *h* could mean 'use', so *h-ear* would be 'use one's ear', i.e. 'hear'. But this is the only pair of words of this kind (there is no **heye* 'use one's eye', **helbow* 'use one's elbow', etc.), and everyone agrees that the resemblances are accidental in this case.

Morphological analysis typically consists of the identification of parts of words, or, more technically, **constituents** of words. We can say that the word *nuts* consists of two constituents: the element *nut* and the element *s*. In accordance with a widespread typographical convention, we will often separate word constituents by a hyphen: *nut-s*. It is often suggested that morphological analysis primarily consists in breaking up words into their parts and establishing the rules that govern the co-occurrence of these parts. The smallest meaningful constituents of words that can be identified are called **morphemes**. In *nut-s*, both *-s* and *nut* are morphemes. Other examples of words consisting of two morphemes would be *break-ing*, *hope-less*, *re-write*, *cheese-board*; words consisting of three morphemes are *re-writ-ing*, *hope-less-ness*, *ear-plug-s*; and so on. Thus, morphology could alternatively be defined as in Definition 2.

Definition 2:
Morphology is the study of the combination of morphemes to yield words.

This definition looks simpler and more concrete than Definition 1. It would make morphology quite similar to syntax, which is usually defined as 'the study of the combination of words to yield sentences'. However, we will see later that Definition 2 does not work in all cases, so we should stick to the somewhat more abstract Definition 1 (see especially Chapters 3 and 4).

In addition to its main sense, where morphology refers to a subdiscipline of linguistics, it is also often used in a closely related sense, to denote a part of the language system. Thus, we can speak of 'the morphology of Spanish' (meaning Spanish word structures) or of 'morphology in the 1980s' (meaning a subdiscipline of linguistics). The term *morphology* shares this ambiguity with other terms such as *syntax*, *phonology* and *grammar*, which may also refer either to a part of the language or to the study of that part of the language. This book is about morphology in both senses. We hope that it will help the reader to understand morphology both as a part of the language system and as a part of linguistics.

One important limitation of the present book should be mentioned right at the beginning: it deals only with spoken languages. Sign languages of course have morphology as well, and the only justification for leaving them out of consideration here is the authors' limited competence. As more and more research is done on sign languages, it can be expected that these

studies will have a major impact on our views of morphology and language structure in general.

1.2 Morphology in different languages

Morphology is not equally prominent in all (spoken) languages. What one language expresses morphologically may be expressed by a separate word or left implicit in another language. For example, English expresses the plural of nouns by means of morphology (*nut/nuts, night/nights*, and so on), but Yoruba uses a separate word for expressing the same meaning. Thus, *ọkùnrin* means '(the) man', and the word *àwọn* can be used to express the plural: *àwọn ọkùnrin* 'the men'. But in many cases where several entities are referred to, this word is not used and plurality is simply left implicit.

Quite generally, we can say that English makes more use of morphology than Yoruba. But there are many languages that make more use of morphology than English. For instance, as we saw in (1.1), Sumerian uses morphology to distinguish between 'he went' and 'I went', and between 'he went' and 'he went to him', where English must use separate words. In Classical Greek, there is a dual form for referring to two items, e.g. *adelphṓ* 'two brothers'. In English it is possible to use the separate word 'two' to render this form, but it is also possible to simply use the plural form and leave the precise number of items implicit.

Linguists sometimes use the terms **analytic** and **synthetic** to describe the degree to which morphology is made use of in a language. Languages like Yoruba, Vietnamese or English, where morphology plays a relatively modest role, are called analytic. Consider the following example sentences.[3]

(1.2) Yoruba

Nwọn	*ó*	*maa*	*gbà*	*pọ́nùn*	*mẹ́wǎ*	*lọ́sọ̀ọ̀sẹ̀*.
they	FUT	PROG	get	pound	ten	weekly

'They will be getting £10 a week.'

(Rowlands 1969: 93)

(1.3) Vietnamese

Hai	*đứ.a*	*bo?*	*nhau*	*là*	*tại*	*gia-đình*	*thàng chông*.
two	individual	leave	each.other	be	because.of	family	guy husband

'They divorced because of his family.'

(Nguyen 1997: 223)

[3] For each example sentence from an unfamiliar language, not only an idiomatic translation is provided, but also a literal ('morpheme-by-morpheme') translation. The key for abbreviations is found on pp. xv–xvi, and further notational conventions are explained in the Appendix to Chapter 2.

When a language has almost no morphology and thus exhibits an extreme degree of analyticity, it is also called **isolating**. Yoruba and Vietnamese, but not English, are usually qualified as isolating. Languages like Sumerian, Swahili or Lezgian, where morphology plays a more important role, would be called synthetic. Let us again look at two example sentences.

(1.4) Swahili
Ndovu wa-wili wa-ki-song-ana zi-umia-zo ni nyika.
elephants PL-two 3PL-SUBORD-jostle-RECP 3SG-hurt-REL is grass
'When two elephants jostle, what is hurt is the grass.'
(Ashton 1947: 114)

(1.5) Lezgian
Marf-adi wiči-n qalin st'al-ra-ldi qaw gata-zwa-j.
rain-ERG self-GEN dense drop-PL-INS roof hit-IMPF-PST
'The rain was hitting the roof with its dense drops.'
(Haspelmath 1993: 140)

When a language has an extraordinary amount of morphology and perhaps many compound words, it is called **polysynthetic**. An example is West Greenlandic.[4]

(1.6) West Greenlandic
Paasi-nngil-luinnar-para ilaa-juma-sutit.
understand-not-completely-1SG.SBJ.3SG.OBJ.IND come-want-2SG.PTCP
'I didn't understand at all that you wanted to come along.'
(Fortescue 1984: 36)

The distinction between analytic and (poly)synthetic languages is not a bipartition or a tripartition, but a continuum, ranging from the most radically isolating to the most highly polysynthetic languages. We can determine the position of a language on this continuum by computing its degree of synthesis, i.e. the ratio of morphemes per word in a random text sample of the language. Table 1.1 gives the degree of synthesis for a small selection of languages.

[4] There is another definition of *polysynthetic* in use among linguists, according to which a language is polysynthetic if single words in the language typically correspond to multi-word sentences in other languages. In this book we will not use the term in this sense, but under such a definition, Swahili would be classified as a polysynthetic language.

Language	Ratio of morphemes per word
West Greenlandic	3.72
Sanskrit	2.59
Swahili	2.55
Old English	2.12
Lezgian	1.93
German	1.92
Modern English	1.68
Vietnamese	1.06

Table 1.1 The degree of synthesis of some languages

Source: based on Greenberg (1959), except for Lezgian

Although English has much more morphology than isolating languages like Yoruba and Vietnamese, it still has a lot less than many other languages. For this reason, it will be necessary to refer extensively to languages other than English in this book.

1.3 The goals of morphological research

Morphological research aims to describe and explain the morphological patterns of human languages. It is useful to distinguish four more specific sub-goals of this endeavour: elegant description, cognitively realistic description, system-external explanation and a restrictive architecture for description.

(i) **Elegant description**. All linguists agree that morphological patterns (just like other linguistic patterns) should be described in an elegant and intuitively satisfactory way. Thus, morphological descriptions should contain a rule saying that English nouns form their plural by adding *-s*, rather than simply listing the plural forms for each noun in the dictionary (*abbot, abbots; ability, abilities; abyss, abysses; accent, accents; …*). In a computer program that simulates human language, it may in fact be more practical to adopt the listing solution, but linguists would find this inelegant. The main criterion for elegance is **generality**. Scientific descriptions should, of course, reflect generalizations in the data and should not merely list all known individual facts. But generalizations can be formulated in various ways, and linguists often disagree in their judgements of what is the most elegant description. It is therefore useful to have a further objective criterion that makes reference to the speakers' knowledge of their language.

(ii) **Cognitively realistic description**. Most linguists would say that their descriptions should not only be elegant and general, but they should