Appendix: The structure of words

1 Introduction

The main part of *A Student's Introduction to English Grammar* concerns syntax – the structure of sentences, and especially clauses. But we often have to mention aspects of the internal structure of words. The endings of certain words often crucially depend on the syntax of the sentence they're in (that's what **inflectional morphology** is about), and occasionally we have had to mention other aspects of the way the basic shapes of lexemes are constructed (**lexical word formation**, also known as **derivational morphology**). In this appendix (the content of which corresponds primarily to Chapters 18 and 19 of *The Cambridge Grammar of the English Language*), we will give a brief (and of course incomplete) overview of the more important features of English word structure.

We'll first draw the distinction between inflection and derivation (§2). Then we'll lay out the theoretical basics of inflectional morphology, with some attention to the differences between speech and writing (§3). Since we'll be concerned much more with written English than spoken English in this appendix, we devote §4 to some very useful rules about how spelling changes when suffixes are added. Then we survey the inflectional morphology of verbs (§5), nouns (§6), and the categories that show grade inflection (§7). A few central points of English lexical morphology then follow in §8.

2 Inflectional morphology and lexical morphology

Morphology deals with the composition and internal structure of words in general, and the way this structure determines the word meaning, rather than the way they combine to make larger units like phrases and clauses. It divides into **inflectional** and **lexical** morphology.

Inflectional morphology deals with the differences between the shapes of the **inflectional forms** of variable lexemes; for example, the formation of the verb-forms *endangers*, *endangered*, and *endangering* from the **lexical base** *endanger*.

Throughout this appendix we use the centred dot '·' to indicate a boundary between word parts when it is relevant: we will write *endanger* as *en·danger*, and refer to the prefix as '*en·*' (to indicate that it joins on at the beginning), and write the plural suffix as '*·s*' (to show that it joins on at the end), and write the middle part of *impregnable* as

'*·pregn·*' to show that it needs other parts on both the beginning and the end to make a word. We'll use this device for both lexical and inflectional boundaries, so we might write *deconstructionists* as *de·con·struct·ion·ist·s*.

Lexical morphology deals with the formation of lexical bases – with the formation of *endanger*, for example, from the prefix *en·* and the morphologically simple base *danger*. This includes the formation of the lexical bases of invariable lexemes, such as *clever·ly*. *Cleverly* doesn't inflect at all (there are no forms **cleverlier* or **cleverliest*), but the fact that it is made up out of *clever* and *·ly* is a fact of lexical morphology.

The motivation for dividing morphology into these two branches becomes clear when we consider two lexemes that are related lexically but belong to different categories:

[1]		LEXEME		INFLECTIONAL FORMS		
	i	***friend*** (N)	*friend*	*friend·s*	*friend·'s*	*friend·s'*
	ii	***friendly*** (Adj)	*friend·ly*	*friend·li·er*	*friend·li·est*	

Inflectional forms are listed horizontally. But the forms in row [i] belong to a different lexeme from those in row [ii]. A dictionary would typically have just two entries to cover the seven different word forms: one for the noun ***friend*** and one for the adjective ***friendly***.

Inflectional morphology deals with the horizontal relationships in [1]: the different shapes that share the **lexical base** of a lexeme. Much of it involves regular patterns. In [i] the shapes are based on the base *friend*; in [ii] the shapes are based on *friendly*.

Lexical morphology deals with the vertical dimension in [1]: the structure of, and the relation between, different lexical bases. In this example, the noun base *friend* is morphologically simple, while the adjective base *friendly* is formed from it by the addition of *·ly*.

The reason we treat the relation between the forms *friendly*, *friendlier*, *friendliest* differently from that between *friendly* and *friend* is that there are **rules of syntax** that determine where the various inflectional forms of a lexeme may or must appear. Suppose, for example, that we want to insert the noun lexeme ***friend*** in the contexts shown in [2i] and the adjective lexeme ***friendly*** in those shown in [2ii]:

[2]					
	i	a.	*She's been a good _____.*	b.	*Their _____ car was a VW.*
	ii	a.	*He's _____ than his brother.*	b.	*He's the _____ of them all.*

If you want to use ***friend*** to fill the blanks in [i], you must use the plain singular form *friend* in [ia], and you must use a genitive form in [ib] – either singular (*their friend's car*) or plural (*their friends' car*). And if you want insert a single word to fill the blanks in [ii] with ***friendly***, you need the comparative form *friendlier* in [iia] and the superlative form *friendliest* in [iib] – any other choices would be ungrammatical (e.g. **He's friendly than his brother*).

These rules apply quite generally: context [ia] requires the plain singular form of any noun you might want to substitute for the blank; [ib] needs a genitive form; [iia] needs a comparative; and [iib] needs a superlative.

By contrast, no rule of syntax says that an adjective appearing in [ii] must be formed from a noun, in the way that the base *friendly* is formed from *friend*. The

blanks in [iia–b] could just as well be filled by *older* and *oldest*, respectively, forms whose lexical base is morphologically simple, not formed from anything else.

Nor is there a rule saying that the lexical base of nouns filling the blanks in [i] must be morphologically simple, like *friend*. We could fill the blank in [ia], for example, with *teacher* (which is derived from the verb *teach* by adding *·er*). It doesn't matter how the lexical base of the lexeme is made up internally, and you don't need to know in order to use it correctly. All that matters is whether you have picked a syntactically admissible inflectional form of the lexeme you decide on.

So **inflectional** morphology ties in mainly with the rules of the grammar, while **lexical** morphology is mainly relevant to the content of the dictionary. Inflection is a matter of variation in the morphological form of lexemes determined wholly or in part by the syntactic context in which the lexeme occurs. Inflectional morphology describes the shape of inflectional forms while syntax describes where they can or must occur. Lexical morphology, on the other hand, relates only to the structure of the words in the dictionary and the formation of new words added to it.

Most of this chapter will concern inflection, in keeping with the focus throughout the book. Our main concern is with the syntax of English, not the dictionary – we're concentrating on how sentences are built rather than how lexemes are structured internally or related to each other.

3 Basics of inflectional morphology

In this section we introduce the basic concepts and terminology needed in the description of English inflection.

3.1 Lexical base

The **lexical base** of a lexeme is the starting point for describing the inflectional forms. In English, the lexical base is almost always identical with one of the inflectional forms. For example, the noun lexeme ***friend*** has the lexical base *friend*, and the plain (i.e. non-genitive) singular form is identical with the lexical base. Likewise, the adjective lexeme ***friendly*** has the lexical base *friendly*, and the plain form is identical with this. The other forms – the plural and genitive forms of the noun, the comparative and superlative forms of the adjective – consist of the lexical base with various suffixes added.

There are a few exceptional lexemes whose lexical base is not identical with any of the inflectional forms. They are lexemes that don't have the full set of inflectional forms normally associated with their category. Specifically, there are a few **plural-only** nouns like *auspices*, *binoculars*, *clothes*, *pliers*, *pants*, *scissors*, and *credentials*. These plurals are formed in the usual way by adding a suffix to the lexical base, but the lexical bases aren't normally found standing alone as a form of the noun lexeme (**an auspice*, **this binocular*, **my clothe*, **the other plier*, . . .).

3.2 Morphological operations

Inflectional forms of a lexeme are formed in various ways, by different **operations** on lexical bases. For example, the plural forms mentioned above are all formed by the operation of **suffixation** of *·s*, i.e., adding *·s* to the end of the base. In English, suffixation is the main operation in the inflectional system, but **modification** of the base also plays an important role. Examples from plural nouns and preterite verbs are shown in [3]:

[3]		PLURAL NOUN FORMATION	PRETERITE VERB FORMATION
i	SUFFIXATION	*dog + ·s = dogs*	*want + ·ed = wanted*
ii	MODIFICATION	*goose* modified = *geese*	*take* modified = *took*

The two operations mentioned in [3] may combine: the plural *wive·s* is formed from *wife* by suffixation of *·s* and also changing the final consonant of the base from *f* to *v*. A few other (relatively minor) operations will be introduced below.

3.3 Shape sharing

As we saw in our discussion of verb inflection in §3.1, the various inflectional forms of a lexeme are not always overtly distinct: two (or more) of them may share the same shape. With ***want***, for example, not only the preterite but also the past participle has the shape *wanted*. The same phenomenon is found with nouns. The plurals of some nouns are identical with the lexical base and hence share the same shape as the singular. With ***bison***, for example, the singular and plural forms share the same shape, *bison*; similarly for ***series***, and others listed in §5.1 below.

However, we never posit two inflectional forms that ALWAYS share the same shape. There must be a distinction in at least one lexeme, for otherwise there would be no morphological difference at all, and hence no basis for drawing an inflectional distinction. The form *written*, for example, occurs in two very different constructions, the perfect (e.g. *He had <u>written</u> the letter*) and the passive (*The letter was finally <u>written</u>*), but there is no verb that has different shapes in these constructions, so from a morphological point of view we have here a single inflectional form, the past participle, not two different forms. The perfect and the passive are SYNTACTICALLY DISTINCT constructions, but they involve THE SAME INFLECTIONAL FORM. Compare also the discussion of the plain form in §3.1.2: no verb has different shapes in imperatives (*<u>Be</u> careful*), subjunctives (*It's essential that he <u>be</u> careful*) and infinitivals (*I will <u>be</u> careful*), so imperative, subjunctive and infinitival are again syntactically distinct constructions containing the same inflectional form.

3.4 Alternation

Very often a given inflectional form is formed in different ways for different subsets of lexemes. For example, while many nouns form their written plural by adding the suffix *·s*, there are others that add *·es*: compare *dog·s* and *fox·es*. This use of the suffixes *·s* and *·es* for the same purpose in different contexts is called an **alternation**. The two shapes are called **alternants**, and the rules of inflectional morphology need to specify the conditions under which one alternant or the other is required.

3.5 The priority of speech

So far in this book we haven't had to talk much about how words or sentences are pronounced. We've just shown them in written form. But when we deal with the internal structure of words, we have to pay some attention to speech, for at least two reasons.

- In the first place, there are alternations in speech that don't show up in writing. The plural suffixes in *cats* and *dogs*, for example, are written the same way but they sound different. Try saying this aloud: *Who let the cats in?* The last part sounds like *sin*. Now say: *Who let the dogs in?* The last syllable sounds more like *zin*. Such differences are never shown in English spelling.
- Secondly, and more importantly for our purposes, the choice between alternants in writing often depends on features of the pronunciation. Consider these two plural nouns:

[4]

	LEXEME	SINGULAR	PLURAL
i	***stomach***	*stomach*	*stomach·s*
ii	***reproach***	*reproach*	*reproach·es*

The reason why we have the *·s* alternant in *stomachs* but *·es* in *coaches* can't be explained by looking at the spelling of the lexical base: both bases end in *ach*. Rather, the alternation in writing reflects the fact that in speech the suffix in *stomachs* is simply a consonant sound, whereas in *coaches* it is made up of a vowel sound plus a consonant sound. And in speech the reason we select the vowel-plus-consonant alternant spelled as *·es* has to do with the way the lexical base is pronounced. This suffix is added to a lexical base that ends in a **sibilant**, a 'hissing sound' like the sounds at the ends of bases like *bliss*, *buzz*, *bush*, *rouge*, *bench*, and *budge*.

Despite the frequent relevance of pronunciation, we don't attempt a full description of inflection in spoken English here. There are several reasons for focusing on writing in a short introductory book like this.

- One is that writing is much more uniform than speech. There are extensive differences in pronunciation between British English and American English, and between these and other regional varieties, whereas differences in spelling are few in number, small in scale, and easy to describe in full.
- A second reason is that the writing system has the advantage of familiarity. Examples in written English can just be shown in their usual written form, whereas exhibiting spoken forms would call for a phonetic alphabet. There is an International Phonetic Alphabet which would serve this purpose, but although it is being used in an increasing number of dictionaries, most people don't know it, and it would take some space to explain.

So what we do here is to continue presenting words and parts of words in ordinary spelling. We'll make informal reference to how they're pronounced when that's necessary to explain the spelling rules.

3.6 Letters and symbols

In describing spelling alternations we need to distinguish between **letters** of the alphabet and **symbols** for sounds. These aren't the same.

- In very simple cases like *hat*, letters and symbols coincide, because each letter happens to be a symbol for a single sound.
- In *heat*, however, there are four letters, though still only three sounds: *ea* is a **composite symbol** representing a single vowel sound.
- Similarly, in *heath* there are still only three sounds, but two composite symbols: *ea* as before, and *th* standing for a single consonant sound.
- And in *sheath* we have six letters, but again only three sounds: *sh*, *ea* and *th* are the composite symbols.

Vowel and **consonant** are terms that by themselves apply purely to speech sounds. Vowels have unimpeded smooth continuing airflow through the mouth, whereas with consonants there is some kind of audible constriction that makes a difference to the sound. When we talk about **vowel symbols** and **consonant symbols**, all we'll mean is symbols representing vowel sounds and symbols representing consonant sounds.

We will have no use at all for the traditional classification of letters into five vowels (*a*, *e*, *i*, *o*, *u*) and twenty-one consonants (all the rest). Take *y*, for example: it's a consonant symbol in *you* and *yacht*; it's a vowel symbol in *by* and *pity*; and in *boy* and *guy* it's neither – it's simply part of a composite symbol.

3.7 Regular and irregular forms

An inflectional form is **regular** if it is formed by a general rule and **irregular** if it is formed by a rule applying only to some fixed number of particular lexemes.

- Take the preterite verb-form *killed*, for example. We say it's a regular form because it is formed by adding *·ed* to the lexical base, like the preterite of most verbs.
- We say that the preterite *drank*, on the other hand, is irregular: the modification of the base vowel here (replacing the vowel heard in *ring* by the vowel heard in *gang*) is found with only a handful of lexemes including ***drink***, ***begin***, ***ring***, ***swim***, etc.
- In some cases, regular and irregular forms co-exist as variants for the same inflectional form. The verb ***burn***, for example, has regular *burned* and irregular *burnt* as variants of the preterite and past participle forms, and a number of other verbs behave similarly (*spelled ~ spelt*, *dreamed ~ dreamt*, etc.).

We call an entire lexeme regular only if ALL its inflectional forms are regular: the general rules must correctly account for every single one of its forms for it to be a regular lexeme. Overall there are around two hundred irregular verbs, some of them very common.

For the most part, forms that are regular in speech are regular in writing, and vice versa. But there are some exceptions in most people's speech, such as those in [5]:

[5]

			IN SPEECH	IN WRITING
i	*says*	3rd sing present tense of ***say***	irregular	regular
ii	*paid*	preterite/past participle of ***pay***	regular	irregular

Says is perfectly regular in writing, but for most speakers it is irregular in speech because it has a different vowel from the lexical base (it rhymes with *fez*, not *faze*). Conversely, *paid* is regular in speech but irregular in writing: the regular form would be **payed* (compare *pray ~ prayed*), but the spelling actually used is different from that.

4 Some spelling rules

In this section we introduce four spelling rules that apply in the formation of more than one inflectional form. They may convince you that although the English spelling system is complex and irregular, it is a little more systematic than you might have realized.

The first three rules actually apply in lexical as well as inflectional morphology. In all of them we need to take account of the pronunciation of bases and suffixes in accordance with the point made above concerning the priority of speech over writing. Keep in mind, though, that relevant features of the spelling system were established several centuries ago and may reflect pronunciations that are no longer current for all or some varieties of English. There are two cases of this kind we should mention here:

- The suffix used to form regular preterites and past participles is now pronounced with a vowel only after bases ending in a *t* or *d* sound, as in *waited* or *landed*: elsewhere it is pronounced as a consonant, as in *scoffed* (which rhymes with *soft*), *feared* (which rhymes with *beard*), etc. In earlier centuries the *·ed* ending was pronounced with a vowel in all cases, and as far as the writing system is concerned it still behaves in all cases as if it were a suffix beginning with a vowel sound.
- In some varieties of English, the sound represented by the letter *r* now occurs only before a vowel: it occurs, for example, in *ram* but not in *mar*, which is pronounced just like *ma*. Most varieties of the English spoken in England are of this kind, whereas most North American varieties still have this sound both before and after vowels. This variation is irrelevant to the writing system, however: bases like *mar* are treated as ending in a consonant sound in all varieties, even the ones where the final *r* has no effect on the pronunciation.

4.1 Consonant-symbol doubling

Consonant-symbol doubling is illustrated in sets of forms like the ones in [6], where *stop* has one *p* but *stopped* has two, and so on:

[6]

i	VERB	*stop*	*stopped*	*stopping*	*stops*
ii	ADJECTIVE	*fat*	*fatter*	*fattest*	
iii	NOUN	*quiz*	*quizzes*		

There is a rule that describes where you get this doubling and where you don't. It can be stated like this:

[7] The final consonant letter of the base is doubled if all of the following conditions are satisfied:
   a. it occurs before a suffix beginning with a vowel sound;
   b. the base ends in a single consonant sound represented by a single letter;
   c. the consonant letter follows a single-letter vowel symbol;
   d. the base is stressed on the final (or only) syllable.

(a) Doubling occurs before suffixes beginning with a vowel

As we noted above, the suffix *-ed* is treated as beginning with a vowel whether or not it is still pronounced with a vowel in Present-day English (*stopped* used to be pronounced more like *stop Ed*, and the spelling rules have largely remained the same since then). So in [6i] we get doubling in *stopped* as well as *stopping*. But there is no doubling in *stops*, where the suffix consists of a consonant. In plural nouns we have doubling in *quizzes*, where the suffix begins with a vowel, but not in *hats*, where again it consists of a consonant.

(b) Base ends in single consonant

There is no doubling in forms like *grasping*, where the base ends in a sequence of two consonants. Nor is there doubling in *sawing*, *boxing*, and the like. It's crucial not to be confused by the spelling: the letter *w* is sometimes a consonant symbol, but the base *saw* is pronounced ending in a vowel, not a consonant. And *box* ends in a sequence of two consonants (*x* represents the two sounds that in the word *phonetics* are represented separately by *c* and *s*). We do have doubling in *marring*, though: as noted above, the base *mar* is treated as ending in a consonant even in those varieties where the final *r* sound is no longer pronounced.

(c) Consonant letter follows single-letter vowel symbol

The consonant symbols doubled in [6] are preceded by the vowel symbols *o*, *a*, and *i* respectively. There is no doubling in forms like *beating*, *roaring*, *cooler*, etc., where the vowel is represented by the composite symbols *ea*, *oa*, *oo*. (The *u* in *quiz* is not part of a composite vowel symbol; it represents a consonant – the one found at the beginning of *wizard*.)

(d) Stress on final (or only) syllable

This condition accounts for the difference in such verbs as ***prefer*** and ***offer***:

[8] a. *prefer* *prefe<u>rr</u>ed* *prefe<u>rr</u>ing* b. *offer* *offe<u>r</u>ed* *offe<u>r</u>ing*

The base *prefer* has the stress on the second syllable, so doubling applies, but *offer* has the stress on the first syllable, which prevents doubling.

The rule given in [7] does most of the work, but English spelling is always quite difficult and, sure enough, there are some exceptions to this last condition.

- Firstly, non-final stress does not prevent doubling of final *l* in British English, as in *travelling* or *crueller* (American English spelling has *traveling* and *crueler*, in keeping with the rule).
- Secondly, there are a number of other verbs with non-final stress where doubling is found, either obligatorily (e.g. *formatted*, *leapfrogged*) or optionally (*benefitted*/*benefited*, *worshipped*/*worshiped*), and those you just have to learn by acquaintance.

4.2 Final *e* deletion

Bases ending in *e* often lose this *e* when a suffix beginning with a vowel is added. There are two different cases, though. The first case is where the *e* is what children are often taught to call the 'magic *e*'; we'll call it '**mute** *e*'. The second case is where the *e* is part of a composite symbol.

[9]	i MUTE *e*	*hope*	*hoping*	*hoped*
	ii PART OF COMPOSITE SYMBOL	*subdue*	*subduing*	*subdued*

(a) Mute *e*

The term **mute** *e* applies to a base-final *e* that is preceded by a consonant symbol and does not itself represent a sound. In speech, for example, the base *hope* ends in a consonant, so the written *e* is mute.

Mute *e* always drops before a suffix beginning with *e*, and normally does so before the *·ing* suffix of gerund-participles. The few verbs where *e* is retained before *·ing* include a number whose base ends in *inge*; the *e* is obligatory in *singeing* (keeping the *e* distinguishes the gerund-participle of ***singe*** from that of ***sing***). In the BrE word ***whinge*** ("complain, grumble", very common in Australian English), the *e* remains optionally: both *whingeing* and *whinging* are found.

(b) Composite vowel symbols ending in *e*

The *e* at the end of *subdue* is part of a composite symbol, *ue*. A final *e* of this kind always drops (like mute *e*) before a suffix beginning with *e*, and sometimes also before *·ing*. It generally drops before *·ing* when the composite symbol is *ue*, but not when it is *ee* (*freeing*), *oe* (*hoeing*), or *ye* (*dyeing*). The composite symbol *ie* is usually replaced by *y* before *·ing*, as in *lying*. Note the contrast between *dying*, a form of ***die*** ("cease living"), and *dyeing*, a form of ***dye*** ("colour with chemicals").

4.3 Final *y* replacement

The third rule applies with bases ending in *y*, normally as a single-letter vowel symbol. Before a suffix we have the alternation shown in [10]:

[10]		TREATMENT OF *y*	CONTEXT			EXAMPLES	
	i	*y* retained	before *·ing* or *·'s*	*deny*	*deny·ing*	*baby*	*baby·'s*
	ii	*y* replaced by *ie*	before *·s*	*deny*	*denie·s*	*baby*	*babie·s*
	iii	*y* replaced by *i*	elsewhere	*deny*	*deni·ed*	*pretty*	*pretti·er*

The replacement of *y* by *ie* occurs in verbs before 3rd person singular present tense *·s* (*denies*), and in nouns before plural *·s* (*babies*), but never before genitive *·'s* (*baby's*). Replacement by *i* is found in verbs before *·ed* (*denied*) and in adjectives before comparative *·er* and superlative *·est* (*prettier*, *prettiest*). ***Dry*** and ***shy*** are exceptions, having *dryer*/*dryest* and *shyer*/*shyest* as optional variants of regular *drier*/*driest* and *shier*/*shiest*. Replacement by *i* is also seen before various suffixes in lexical morphology (as in *denial*, *embodiment*, etc.).

As we noted above, these rules normally apply only when *y* represents a vowel by itself; when *y* is part of a composite symbol, it is retained in all contexts, as in *buys*, *boys*, *played*, *coyer*, etc., except for a few irregular forms where *ay* is replaced by *ai*, as in *lain*, *paid*, *said*.

4.4 Alternation between ·*s* and ·*es*

This alternation occurs with the 3rd person singular present tense suffix in verbs and the plural suffix in nouns. In speech, the corresponding alternation is also found with the genitive suffix in singular nouns but in writing this is invariably *·'s* (in regular forms). In speech it is found not only in plural *cats* and *foxes* but also in genitive singular *cat's* and *fox's*, which sound exactly the same as the plurals. In writing things are quite different: the genitives have the same suffix, and it is different from the plural one.

The alternation between *·s* and *·es* can best be described by taking *·s* as the default form of the suffix and stating where it is that you have to use *·es* instead. There are two cases to consider.

(a) The *·es* alternant represents spoken vowel + consonant

The *·es* alternant is added to bases which in speech take a suffix with the form of vowel + consonant. As noted above, these are bases that end in a sibilant, or hissing sound:

[11]							
	i	*kiss*	*rose*	*bush*	*rouge*	*bench*	*judge*
	ii	*kiss·es*	*ros·es*	*bush·es*	*roug·es*	*bench·es*	*judg·es*

A good number of bases which in speech end in a sibilant end in writing with mute *e*: in the examples in [11] this applies to *rose*, *rouge*, and *judge*. However, this mute *e* drops by the general rule of *e* deletion described in §3.2 above, and the suffixed forms in [ii] can be treated in a uniform way.

(b) The *·es* alternant is commonly required after bases ending in consonant + *o*

Bases ending in *o* commonly take *·es* if the *o* follows a consonant symbol; otherwise they take the default *·s* alternant. This rule is illustrated in [12]:

[12] i When the *o* follows a consonant symbol, we typically get *·es*:
go·es hero·es potato·es tomato·es torpedo·es veto·es
ii When the *o* does not follow a consonant symbol, we invariably get *·s*:
boo·s embryo·s folio·s radio·s video·s zoo·s

There is some difference between verbs and nouns here. With verbs, the *·es* alternant is almost invariably used when *o* follows a consonant – as in the verb uses of *goes*, *torpedoes* and *vetoes*. The same applies to nouns which are identical in form to verbs (*goes*, *torpedoes*, *vetoes*, *echoes* and *embargoes* can be either 3rd singular present tense verbs or plural nouns). But a good number of other nouns with *o* following a consonant take the default *·s* instead: this is obligatory for *dynamo*, *kilo*, *piano*, *Eskimo*, and optional with *bongo*, *buffalo*, *halo*, *motto*, *volcano*, and some others.

5 Verb inflection

The inflectional categories of verbs were introduced in §3.1, where we discussed their meaning and syntactic distribution. In this chapter we're concerned solely with their morphological formation – which forms have which suffixes or modifications, and how the results are spelled. Our main focus will be on **lexical** verbs (i.e. verbs other than auxiliaries).

Almost all lexical verbs have six inflectional forms. The plain form and the plain present tense are identical with the lexical base, so we don't need to say any more about them here (recall that it is just with the auxiliary verb ***be*** that the plain form is distinct from all present tense forms). Of the others, the gerund-participle and the 3rd person singular present tense are very straightforward, so we'll deal with them first. Then we'll turn to the preterite and past participle forms, where we find virtually all of the considerable complexity in English verb inflection.

5.1 The gerund-participle

The gerund-participle is invariably formed by adding the suffix *·ing* to the lexical base. In speech, that is all there is to it; even *be·ing* is completely regular. In writing, addition of the suffix may lead to modification of the base involving consonant doubling, *e* deletion and replacement of *ie* by *y*, as described in §§3.1–3.3:

[13] i LEXICAL BASE *see stop hope subdue hoe lie*
ii GERUND-PARTICIPLE *see·ing stopp·ing hop·ing subdu·ing hoe·ing ly·ing*

5.2 The 3rd person singular present tense

This is normally formed by adding *·s* or *·es* to the base. But in this case ***be*** is irregular: we get *is*, not **bes*. ***Have*** is also irregular, losing the *ve* of the base: *has*, not **haves*. In speech, *does* is also irregular in that the vowel differs from that of the base (it rhymes with *buzz*, not *booze*). The same applies to *says* (as we noted in connection with [5i], for most speakers it rhymes with *fez*, not *faze*).

The choice between *·es* and the default alternant *·s* has been described in §3.4. A sample of forms is given in [14]:

[14]								
	i	LEXICAL BASE	*miss*	*lose*	*touch*	*go*	*boo*	*stop*
	ii	3RD SING PRESENT	*miss·es*	*los·es*	*touch·es*	*go·es*	*boo·s*	*stop·s*

5.3 The preterite and past participle

We take these two forms together, since for all regular verbs and a high proportion of irregular ones they are morphologically identical – they have exactly the same shape.

Regular forms

We begin with verbs whose preterite and past participle are regular in speech. In writing, these are formed by the addition of the suffix *·ed*, with consonant doubling, *e* deletion and replacement of *y* by *i* applying as described in §§3.1–3.3. Examples are given in [15]:

[15]								
	i	BASE	*laugh*	*stop*	*prefer*	*hope*	*subdue*	*deny*
	ii	PRET.	*laughed*	*stopp·ed*	*preferr·ed*	*hop·ed*	*subdu·ed*	*deni·ed*
	iii	PST PRT	*laughed*	*stopp·ed*	*preferr·ed*	*hop·ed*	*subdu·ed*	*deni·ed*

Irregular forms with preterite and past participle identical

Many irregular verbs are like regular ones in having shape sharing between preterite and past participle. There are a dozen or so where there is variation between a regular form and a mildly irregular one, as with ***spell***, whose preterite can be *spelled* or *spelt*.

We find a considerable range of morphological relations between the preterite and past participle and the lexical base. A sample of these are illustrated in [16], with commentary given below:

[16]			A	B	C	D	E	F	G	H
	i	BASE	*hit*	*bend*	*burn*	*keep*	*flee*	*think*	*dig*	*stand*
	ii	PRET	*hit*	*bent*	*burnt*	*kept*	*fled*	*thought*	*dug*	*stood*
	iii	PST PRT	*hit*	*bent*	*burnt*	*kept*	*fled*	*thought*	*dug*	*stood*

Type a. Preterite and past participle identical with base

There are over twenty verbs where the preterite and past participle are both identical with the lexical base. Most of them have bases ending in *t*, but there are also a few in

d, such as *shed*. Some lexemes, such as ***quit***, have variant regular forms: *He quit* or *He quitted*.

Type b. Base-final *d* replaced by *t*

Similar to Type A are those where the preterite and past participle are not identical to the base but differ simply in the replacement of final *d* by *t*. They include ***build***, ***send***, ***spend***, etc.

Type c. Preterite and past participle formed by addition of *·t*

This is a small class with all members having regular variants too: *burnt* or *burned*. Bases ending in *ll* lose one *l* before *·t*: *smell ~ smelt*. Other examples include ***dwell***, ***learn***, ***spoil***

Type d. Addition of *·t* with modification of the base

These are similar to Type C, but the addition of *·t* is accompanied by modification of the base. The modification is usually just a change in the vowel, but a more extensive modification is found in *leave ~ left*. Other members of the class include ***feel***, ***mean***, ***sleep***. ***Leap*** is Type D in speech, but Type C in writing.

Type e. Addition of *·d* with modification of the base

A small number of verbs add *·d* rather than *·t*, in all cases with modification of the base. Other examples include ***hear***, ***say***, ***sell***; with *said* the written form is one of the exceptional cases of replacement *y* by *i* within the composite symbol *ay*.

Type f. Preterites and past participles in *ought* or *aught*

A handful of verbs have highly irregular preterites and past participles with *ought* or *aught* (which are pronounced alike) replacing vowel + any following consonants in the base. Other examples include ***buy***, ***catch***, ***seek***. They might be subsumed under Type D, but the final *t* can hardly be analysed as a suffix.

Type g. Preterite and past participles formed by vowel change

In a fair number of verbs the preterite and past participle differ from the base just in respect of the vowel. A variety of different vowel pairs are found: compare *hang ~ hung*, *find ~ found*, *shine ~ shone*, etc.

Type h. Miscellaneous

There are a few verbs that don't fit into any of the above patterns, and where the preterite and past participle differ from the base in a unique way. *Stood* differs from *stand* in the vowel and loss of *n*; *had* and *made* differ from *have* and *make* in the second consonant.

Irregular forms with preterite and past participle distinct

We turn finally to verbs where the preterite and past participle are different in shape. In most cases the past participle contains a distinctive suffix spelled in the three alternant ways shown in [17] and added to the base or, in some cases (marked here by underlining), to the preterite form:

[17]								
	i	*·n*	after vowel symbol or *re*:	*grown*	*lain*	*seen*	*sewn*	*torn*
	ii	*·ne*	for ***bear***, ***do***, and ***go***:	*borne*	*done*	*gone*		
	iii	*·en*	elsewhere:	*broken*	*chosen*	*fallen*	*swollen*	*taken*

- *Grow*, *see*, and *sew* end in composite vowel symbols, and *·n* is added directly to them. With *lay* we have an exceptional case of *y* replacement applying to the composite symbol *ay*. When the suffix is added to the preterite form *tore* the mute *e* following *r* is deleted (and the same applies with *worn*, formed from *wore*, the preterite of ***wear***).
- The items in [ii] are simply exceptions; there is no general rule that assigns them the *ne* spelling. All have a different vowel sound than the base.
- The *·en* alternant is found in all other cases. When it is added to a form ending in mute *e*, the ordinary rule of *e* deletion applies, as we see in *broken*, *chosen*, and *taken*. In *swollen* the vowel of the base *swell* is changed.

Again we find a variety of morphological relations between lexical base, preterite and past participle. Here's a sample of examples, with discussion following.

[18]		A	B	C	D	E	F
i	BASE	*show*	*take*	*ride*	*lie*	*drink*	*fly*
ii	PRETERITE	*showed*	*took*	*rode*	*lay*	*drank*	*flew*
iii	PAST PARTICIPLE	*shown*	*taken*	*ridden*	*lain*	*drunk*	*flown*

Type a. Preterite: regular; past participle: base + suffix

Showed is a regular preterite, while *shown* consists of the base and the distinctive past participle suffix. Other such verbs include ***mow*** and ***sew***, which, like ***show***, have bases ending in a vowel sound and hence take the *·n* alternant; ***prove*** ends in a consonant sound and hence takes the default *·en* alternant. All these verbs have a regular past participle as a variant of the irregular one.

Type b. Preterite: vowel change; past participle: base + suffix

Here the past participle is formed in the same way as in Type A, but the preterite is irregular, formed by vowel change. Other examples: ***blow***, ***eat***, ***give***.

Type c. Preterite: vowel change; past participle: modified base + suffix

These verbs differ from those of Type B in that the base to which the past participle suffix is added is modified. In speech the base vowel is changed; in writing mute *e* is

dropped and *d* or *t* is doubled. Other examples: ***smite***, ***stride***, ***write***. Such verbs as ***drive*** and ***rise*** belong here in speech but in Type B in writing.

Type d. Preterite: vowel change; past participle: preterite form + suffix

In this type the past participle suffix is added not to the lexical base but to the preterite form. Other examples: ***tear*** and ***wear*** (mentioned above), and ***break***, ***choose***, ***tread*** (with the general rule of consonant doubling applying with *trod* to give *trodden* for the past participle). One of the small differences in verb inflection between AmE and BrE is that ***get*** belongs in this class in AmE (with the three forms *get*, *got*, *gotten*), whereas in BrE the past participle has the same shape as the preterite, *got*, which means ***get*** belongs in Type G of [16].

Type e. Three different vowels; no suffix

With a few verbs the three forms are distinguished solely by their vowels. Other examples include ***begin***, ***ring***, ***swim***. All have the same pattern of vowel change, with *i* ~ *a* ~ *u* as the written vowel symbols

Type f. Miscellaneous

There are a few verbs which don't fit into any of the above patterns. With ***fly***, the three forms have different vowels but the past participle also contains the suffix; similarly with ***do***.

With ***be*** and ***go*** the preterite bears no resemblance at all to the lexical base, while the past participle consists of base + suffix. With ***come*** and ***run*** the past participle is identical with the lexical base but the preterite is formed by vowel change.

6 Noun inflection

Nouns inflect for **number** and for **case**. The non-genitive singular, or **plain singular**, is identical with the lexical base. What we need to consider here is the marking of **plural** number and **genitive** case. (We don't cover the inflection of pronouns or the determinatives ***this*** and ***that*** here; the irregular forms are listed in Ch. 5, and there's nothing we can add here about how they're formed.)

6.1 Plural formation

Plurals which are regular in speech are formed in writing by adding the default *·s* or else *·es* to the lexical base. The choice between these alternants has been discussed in §3.4 and is illustrated again in [19]:

[19]

i	LEXICAL BASE	*cross*	*horse*	*edge*	*echo*	*book*	*studio*
ii	PLURAL	*cross·es*	*hors·es*	*edg·es*	*echo·es*	*book·s*	*studio·s*

With *horse* and *edge*, addition of *·es* triggers *e*-deletion in the base.

Plurals that are irregular in speech we discuss under four headings.

(a) Modification of the base-final consonant

With a good number of nouns, addition of the plural suffix is accompanied by a modification of the consonant at the end of the base. When the consonant in question is represented in writing by *f*, the modification is reflected in the spelling, as in:

[20]								
	i BASE	*calf*	*knife*	*leaf*	*loaf*	*thief*	*wife*	*wolf*
	ii PLURAL	*calve·s*	*knive·s*	*leave·s*	*loave·s*	*thieve·s*	*wive·s*	*wolve·s*

The consonant symbol *f* is changed to *v*, and mute *e* is added if not already present. The default is for noun bases not to undergo such modification: words such as *belief*, *chief*, *proof*, *safe* don't, and when an invented word ending in *f* is added to the lexicon, its plural does not show the modification (the noun *Smurf* appeared in English with a TV cartoon series in the 1980s, and the plural is of course *Smurfs*).

Some words, such as *dwarf* and *hoof*, have both regular and irregular plurals: *dwarfs* and *hoofs* or *dwarves* and *hooves*.

In speech the noun *house* undergoes a similar modification, and so do some nouns ending in the consonant represented as *th* (e.g. *mouth*), but this is not reflected in the spelling.

(b) Vowel change and the suffix *·en*/*·ren*

With a small number of nouns the plural is formed by changing the vowel and/or adding the suffix *·en* or *·ren*. Examples are given in [21]:

[20]								
	i BASE	*foot*	*tooth*	*mouse*	*man*	*woman*	*ox*	*child*
	ii PLURAL	*feet*	*teeth*	*mice*	*men*	*women*	*ox·en*	*child·ren*

With *woman*, both vowels are changed in speech (*women* rhymes with *him in*), but only the second in writing. With *mouse* (and also *louse*) the vowel change is accompanied by a change in the consonant symbol. In speech, *children* shows both a suffix and a vowel change in the base (notice that *child* rhymes with *filed*, but the beginning of *children* sounds like *chill*).

(c) Base plurals

A fair number of nouns have plurals that are, like the singular, identical with the base:

[22]								
	BASE = PLURAL	*sheep*	*cod*	*bison*	*barracks*	*series*	*Chinese*	*Roma*

Most of these nouns belong to one or other of the following categories:

- Nouns denoting edible fish and game animals – creatures that are traditionally hunted. *Cod* and *bison* from [22] belong here; others include *salmon*, *trout*, *deer*, *grouse*, *reindeer*. Some (such as *elk*) have a regular plural alternant. For animals that have never been the target of hunting or fishing, the base plural is impossible: **three cockroach*, **several spider*, **two large dog*.

- Nouns with bases ending in *s* (a single *s*, not double): *barracks*, *headquarters*, *means*, *series*, *species*, etc. (we never find **three different barrackses*).
- Nationality nouns ending in *·ese*: *Chinese*, *Japanese*, *Vietnamese*, etc. (we never find **millions of Chineses*).
- Many names of ethnic groups: *Apache*, *Bedouin*, *Inuit*, *Kikuyu*, *Navajo*, *Roma*, etc. With most of these a regular plural alternant is sometimes used, but not as commonly as the base plural.

(d) Foreign plurals

A considerable number of nouns of Latin, Greek, and various other origins have plurals taken from those languages. Many belong to scientific or otherwise relatively learned vocabulary. A good proportion have regular plurals as variants, and these tend to be preferred in informal contexts. Some examples are given in [23]:

[23]						
	i BASE	*formula*	*larva*	*stimulus*	*syllabus*	*phenomenon*
	ii FOREIGN PLURAL	*formulae*	*larvae*	*stimuli*	*syllabi*	*phenomena*
	iii REGULAR PLURAL	*formulas*	–	–	*syllabuses*	–

With some of these nouns the foreign plural is much more common than the singular, and for some speakers has been learned as a non-count singular. Examples include *data* (from *datum*), *media* (*medium*), *algae* (*alga*), *bacteria* (*bacterium*), *criteria* (*criterion*), and *phenomena* (*phenomenon*). There is considerable variation, however, with respect to how far this singular usage has become established as a variant of the plural. Singular *data* and *media* are firmly established in Standard English (examples like *The data is reliable* or *The media keeps hounding her* are common and widely considered acceptable); but singular *criteria* and *phenomena* are not (expressions like !*this criteria* or !*a new phenomena* are regarded as non-standard). Various other items, such as *algae* and *bacteria*, belong in the middle ground.

6.2 Genitive formation

The genitive forms of the personal pronouns and the interrogative/relative genitive pronoun *whose* have been described in §5.8.3, and §11.3, while general issues relating to the genitive construction are discussed in §5.9: in this chapter we confine our attention to the two kinds of genitive illustrated in [24], the *'s* **genitive** and the **bare genitive**:

[24]						
	i *'s* GENITIVE	*girl ~ girl's*	*woman ~ woman's*	*women ~ women's*	*James ~ James's*	
	ii BARE GENITIVE	*girl ~ girls'*	*barracks ~ barracks'*		*James ~ James'*	

The *'s* genitive is the default form: we need only give the particular conditions under which the bare genitive must or may occur and can then say that the *'s* genitive is used for the remainder.

(a) The bare genitive

In speech, the bare genitive is not overtly marked at all, being identical in form with its non-genitive counterpart. In writing, it is marked by a final apostrophe.

The bare genitive is virtually restricted to nouns ending in *s*. It is usually obligatory, but it may also be optional, alternating with the *'s* genitive:

[25]				
i PLURALS FORMED WITH THE *·s/·es* SUFFIX	obligatory	*girls'*	*foxes'*
ii NOUNS ENDING IN *s* WITH BASE PLURALS	obligatory	*barracks'*	*series'*
iii CERTAIN PROPER NOUNS ENDING IN *s*	optional	*James'*	*Socrates'*

- The obligatory bare genitives are found with plural nouns marked as such by the *·s* or *·es* suffix (including those with modification of the lexical base, as in *wives*'), and with nouns ending in *s* that have base plurals, as described in §5.1 above. Note that *barracks'* and *series'*, like the non-genitive counterparts, can be either singular or plural.
- The optional bare genitive is found in proper nouns ending in a single *s*, especially classical ones. It is more likely in writing than in speech, and more formal than the variant with *'s*.

(b) The *'s* genitive

In writing, the *'s* genitive is invariably formed by adding *'s* to the non-genitive counterpart – which may be a singular (*woman's*) or a plural that is not marked by the *·s/·es* suffix (*women's*). In speech it has the same form and alternation between *·s* and *·es* as the regular plural suffix: *fox's*, for example, is pronounced the same as *foxes*.

The genitive *'s* sounds exactly like the plural suffix *·s/·es*, but there is an interesting difference: the genitive does not trigger any modification of the base. The genitive of *wife* is *wife's*, not **wive's* – and in speech *mouth's* is pronounced differently from *mouths*.

7 Grade inflection

The last of the three inflectional systems of English to consider is that of **grade**, with three contrasting terms: plain, comparative, and superlative. It applies primarily to adjectives but is found also with a few lexemes from other categories, most clearly adverbs and determinatives. Regular forms are illustrated in [26]:

[26]	ADJECTIVE				ADVERB	DETERMINATIVE
i PLAIN	*cold*	*hot*	*rare*	*easy*	*soon*	*few*
ii COMPARATIVE	*cold·er*	*hott·er*	*rar·er*	*easi·er*	*soon·er*	*few·er*
iii SUPERLATIVE	*cold·est*	*hott·est*	*rar·est*	*easi·est*	*soon·est*	*few·est*

The plain form is identical with the lexical base while the comparative and superlative forms are marked by the suffixes *·er* and *·est*. These begin with a vowel, which triggers the modification of the base by the general spelling rules given in §§3.1–3.3 above: consonant doubling with *hot*, *e* deletion with *rare*, *y* replacement with *easy*.

There are a few lexemes where the comparative and superlative forms are highly irregular, bearing little if any resemblance to the plain form. These include the following:

[27]					
	i PLAIN	*good/well*	*bad/badly*	*much/many*	*little*
	ii COMPARATIVE	*better*	*worse*	*more*	*less*
	iii SUPERLATIVE	*best*	*worst*	*most*	*least*

- ***Good*** and ***bad*** are adjectives, while ***well*** and ***badly*** can be either adjectives (as in *I'm feeling well*/*badly*) or adverbs (*He behaved well*/*badly*).
- ***Much***, ***many***, and ***little*** are determinatives. The determinative ***little*** (as in *It has little merit*) is a different lexeme from the adjective ***little*** (as in *a little creature*), which we're not referring to here: the adjective has regular inflection (though its inflected forms are now rare).

Inflectional and non-inflectional marking of grade

The comparative and superlative categories differ from those discussed in §§4–5 above in that they can be marked by the separate words *more* and *most* as well as by means of inflection. Some lexemes have only inflectional comparatives and superlatives, others have only the non-inflectional type, while others accept both. These comparative examples illustrate:

[28]	TYPE	INFLECTIONAL COMPARATIVE	NON-INFLECTIONAL COMPARATIVE
	i INFLECTIONAL	*This is better than that.*	**This is more good than that.*
	ii NON-INFLECTIONAL	**This is usefuller than that.*	*This is more useful than that.*
	iii EITHER	*This is gentler than that.*	*This is more gentle than that.*

- Very few lexemes accept only the inflectional type. The clearest examples, besides *good*, are the determinative ***few*** from [26] and the irregular lexemes given in [27].
- The *·ly* suffix is never compatible with grade inflection: neither **clearlier* nor **clearerly* is possible. This means that grade for *·ly* adverbs must always be expressed in the non-inflectional way (*You should speak more clearly*).
- Lexemes with monosyllabic bases nearly always prefer the inflectional type, but with some, such as *fake*, *ill*, *real*, *right*, and *wrong*, it is the other way round, with the inflected forms being very rare.
- Adjectives with two-syllable bases accept non-inflectional grade marking, but inflected forms are often not available, especially for the ones that do not have final stress. Bases of this kind formed with *·y* do take inflection (*sticky*, *stickier*, *stickiest*), but examples of adjectives that do not accept inflection include *brutish*, *careful*, *legal*, *jealous*, *public*, and others formed with the suffixes *·ish*, *·ful*, *·al*, *·ous*, and *·ic*.

8 Lexical morphology

Lexical morphology is concerned with the formation and structure of the lexical bases of lexemes. It is COMPLEMENTARY with inflectional morphology: it deals with those aspects of the formation and structure of words that are not a matter of inflection.

8.1 The structure of words

Bases and affixes

The two main kinds of morphological unit are **bases** and **affixes**. As a starting-point, we can distinguish between them as follows:

[29] i BASE usually a **free** element, one able to stand alone as a word, e.g. *cat*, *for*, *happy*, *luck*, *often*, *resent*, *some*, *they*, *usual*

ii AFFIX a **bound** element, one unable to stand alone as a word e.g. *·ation*, *dis·*, *en·*, *·ish*, *·ist*, *·ly*, *·ment*, *·ness*, *un·*, *·y*

Affixes always attach to a base and are subdivided into **prefixes** and **suffixes**. Prefixes precede the base, as in *un·happy*; suffixes follow the base, as in *resent·ment*. When citing affixes, we insert the morphological boundary symbol before suffixes and after prefixes in order to show that they are bound, always occurring as part of some larger unit rather than as a word on their own.

The qualification 'usually' in [29i] is needed because there are exceptional bases which cannot stand alone as words. Two kinds of **bound base** are illustrated in *dur·able* and *scissor·like.*

- The word *durable* was borrowed as a whole from French, not created by the operations of word formation in English – and in French the base, spelled with mute *e* as *dure*, is free. The other component of the word, *·able*, is recognizable as the suffix that combines with free bases in innumerable words like *enjoyable*, *perishable*, *readable*, *retrievable*. Although it is bound in English, *dur·* occupies the same place in word structure as free bases like *enjoy*, *perish*, *read*, *retrieve*, etc.
- As we noted in §2, there are a number of plural-only nouns which have no singular forms. A base like *scissor·* is bound, occurring in *scissors* and certain derived words, but not on its own.

Individual affixes typically attach to a good number of bases, whereas bases usually combine with just a small number of affixes. Of the affixes mentioned above, for example, *un·*, *·ly* and *·able* occur with innumerable bases, and *·y* with a considerable number, whereas the base *luck* combines with just *·y* and *·less* (*lucky*, *luckless*), *usual* with *un·*, *·ly*, *·ness* (*unusual*, *usually*, *usualness*), and so on.

The layered structure of words

One base may contain a smaller one within it. For example, *unhappy* is a base (it can stand alone as a word) consisting of the prefix *un·* and the base *happy*. With *gentleman* we have a base consisting of two smaller bases, *gentle* and *man*. Recall that in our

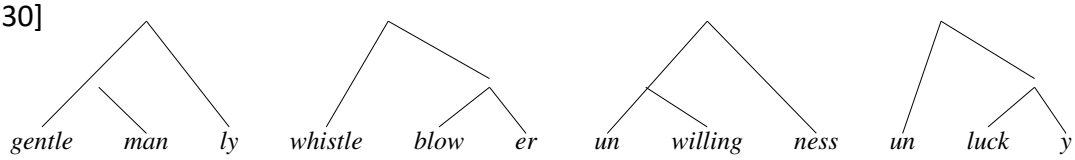
discussion of inflectional morphology we used the term **lexical base** as the unit that is the starting point for the rules determining the shape of the various inflectional forms of a lexeme. The example we gave at the beginning of the chapter was *endanger*, consisting of the prefix *en·* and the base *danger*: *endanger* is the lexical base from which the 3rd person singular present tense *endangers*, the gerund-participle *endangering* and the preterite and past participle *endangered* are formed.

Bases containing more than two elements almost always have a layered structure defining a hierarchy, very similar to the hierarchical structure we see in syntax. For example:

- In the adjective *gentlemanly* the bases *gentle* and *man* combine to form *gentleman*, and the suffix *·ly* is added to that: it's not *gentle* + *manly*, it's *gentleman* + *·ly*.
- In the noun *whistleblower*, on the other hand, the suffix *·er* is added to *blow* to form the base *blower*, and *whistle* combines with this: it's not *whistleblow* + *·er*, it's *whistle* + *blower*.

A similar difference in the hierarchical structure is seen in such a pair as *unhappiness* (with *un·* added to *happy*, forming one adjective base from another, and *·ness* added to *unhappy*, forming a noun base from an adjective base) and *un·luck·y* (with *·y* added to *luck*, forming an adjective base from a noun base and *un·* added to *lucky*, again forming one adjective base from another). Structural diagrams for the four examples are shown in [30]:

[30]



8.2 Morphological operations

There is more to lexical morphology than attaching affixes to bases or adding bases together, though these are certainly the most important ways of forming lexical bases. We need to allow for a fairly wide range of **morphological operations** that are involved in the formation of lexical bases. Indeed, one of the main reasons why grammars are divided into two components, syntax and morphology, is that the ways words are formed differ in significant ways from the way words combine to form sentences.

In the following brief survey, we begin with the two operations that are a matter of combining bases or bases and affixes, and then introduce others that go beyond this elementary pattern.

(a) Compounding

Compound bases are formed from a combination of smaller bases – almost always two. We illustrate here with compound nouns, adjectives and verbs:

[31]								
	i	NOUN	*birdcage*	*gentleman*	*hangman*	*loudmouth*	*outpatient*	*stage-manager*
	ii	ADJ	*dirt-cheap*	*heart-breaking*	*heart-broken*	*skin-deep*	*snow-white*	*stress-free*
	iii	VERB	*baby-sit*	*blow-dry*	*gift-wrap*	*over-react*	*sleepwalk*	*underachieve*

- Compound nouns constitute by far the largest and most varied category. Most denote a subset of what is denoted by the second component: a birdcage is a kind of cage, a gentleman is a kind of man, and so on. But there are certainly a good number that do not have this interpretation: a loudmouth is not a kind of mouth (but a person with a loud mouth – one who talks a lot, typically in an obnoxious way). The second base is most often a noun, while the first can belong to a range of categories: *bird* is a noun, *gentle* an adjective, *hang* a verb, *out* a preposition.
- Compound adjectives often similarly have a denotation included in that of the second element: if something is dirt-cheap, then it must be cheap, and if it's snow-white, it's white. But there are numerous cases where this is not so. A stress-free job is not a free job (it's free of stress), and something that is only skin-deep is not deep. *Heart-breaking* and *heart-broken* illustrate a frequent type where the first base is a noun and the second a gerund-participle or past participle form of a verb.
- There are many compound verbs with a preposition such as *over*, *under*, *out* as first base, but for the rest compound verbs are far less numerous than compound nouns and adjectives.
- One distinctive type of compound noun combines two normally bound bases taken from Greek or Latin, as in: *osteometry*, *osteopath*, *psychopath*, *pathology*,

psychology, etc. Such 'neo-classical' compounds, as they are commonly called, figure very prominently in the learned and scientific vocabulary of the language.

(b) Affixation

In affixation a base is expanded by the addition of a **prefix** at the beginning of the base or a **suffix** at the end. Very often the effect of affixation is to change the primary category of the base – to form a noun from an adjective or verb, to form a verb from a noun or adjective, and so on. We speak here of **category-changing** affixes, as opposed to **category-preserving** affixes:

[32]

		CATEGORY-CHANGING		CATEGORY-PRESERVING	
i	PREFIX	*be·friend*	*en·danger*	*un·happy*	*re·open*
ii	SUFFIX	*wet·ness*	*achiev·able*	*green·ish*	*lion·ess*

Befriend and *endanger* are verbs formed from nouns, while *wetness* is a noun formed from an adjective, and *achievable* is an adjective formed from a verb. Most category-changing affixes are **suffixes**.

Some affixes can be used in EITHER a category-changing OR a category-preserving way.

- The suffix *·ly*, for example, often derives adverbs from adjectives (*rapid·ly*) and, much less often, derives adjectives from nouns (*friend·ly*, *gentleman·ly*, *prince·ly*); in both cases it is category-changing.
- In a few lexemes, however, it derives adjectives from more elementary adjectives. These are not very common in the contemporary language, but examples include *good·ly* "considerable", *kind·ly* "benevolent", and *poor·ly* "not in good health". (*Kindly* and *poorly* also exist as adverbs, of course; but there's a clear grammatical way to distinguish them: grade inflection never occurs with *·ly* adverbs, so although the adjective ***kindly*** has comparative *kindlier* and superlative *kindliest*, those can't be used as adverbs.)

Affixation is commonly accompanied by **modification** of the base, sometimes just in spelling, and sometimes in pronunciation as well. In *achievable*, for example, the mute *e* of *achieve* is dropped, while in *persuasion* we have a change in the consonant at the end of *persuade*.

(c) Conversion

Whereas the verb *hospitalize* is formed from *hospital* by adding the suffix *·ise*, the verb *bottle* (as in *Where do they bottle Coca-Cola?*) is formed from the noun *bottle* without any change of shape at all. This is called **conversion**: a base of one category is formed by extending the use of a base of another category. This is much the most important of the morphological processses that go beyond just adding bases together or attaching affixes to bases.

The main types of conversion are illustrated in [33]:

[33]

		PRIMARY USE	CONVERSION
i	NOUN TO VERB	*needing* <u>*water*</u>$_{N}$	<u>*water*</u>$_{V}$ *the plants*
ii	VERB TO NOUN	<u>*try*</u>$_{V}$ *to persuade her*	*a good* <u>*try*</u>$_{N}$
iii	ADJECTIVE TO NOUN	*very* <u>*professional*</u>$_{ADJ}$	*like a* <u>*professional*</u>$_{N}$
iv	NOUN TO ADJECTIVE	*the* <u>*key*</u>$_{N}$ *to her success*	*a* <u>*key*</u>$_{ADJ}$ *news item*
v	ADJECTIVE TO VERB	*bottles that are* <u>*empty*</u>$_{ADJ}$	<u>*empty*</u>$_{V}$ *the bottle*
vi	VERB TO ADJECTIVE	*film that is* <u>*boring*</u>$_{V}$ *us*	*a* <u>*boring*</u>$_{ADJ}$ *film*

Type [v] differs from the others in that it is not the plain form, or lexical base, of the verb that undergoes conversion, but an inflected form. In this example it is a gerund-participle, but it can also be a past participle as in *It had* <u>*bored*</u>$_{V}$ *them*, which gives via conversion *They were very* <u>*bored*</u>$_{ADJ}$.

There is nothing in the base created by conversion to mark it as such. The direction of conversion reflects the distinction between primary and extended SENSES:

- the primary sense of *bottle* is to denote a narrow-necked container for liquids, and the verb *bottle* incorporates that sense: it means "put into a bottle";
- the primary sense of *water* is that it denotes the physical substance H_2O, and the verb *water* incorporates this sense: it means "provide with water".

There are considerably fewer clear and well-established examples of type [iv], noun to adjective conversion, than of the other types. Both nouns and adjectives can function as heads of attributive modifier in nominals and as head of a predicative complement, so the boundary between the categories is not always clear-cut. Some grammars, for instance, treat material nouns like *cotton* as adjectives in examples like *This shirt is cotton*, but the fact that we can add an adjective as modifier, as in *This shirt is pure cotton*, argues against a conversion analysis here. *Key* in [33] is a clear example because the noun *key* is a count noun and hence takes a determiner in the singular, as in our example *This was the key to her success*, while the absence of a determiner in *This was key to further progress* shows that there has been conversion to an adjective. As we pointed out in §2.5.9, this is an area where there is still ongoing syntactic change and hence divided usage, with conservative speakers still treating *fun*, say, as just a noun, whereas for others it has acquired adjectival properties. We find *very fun*, and *so fun*, *how fun*, and modification by other degree adverbs, and some speakers use inflection for comparative and superlative grade (*funner* and *funnest*). Other examples of words currently undergoing such extensions of use include *core* (like *key*), *bandaid*, *draft*, *freak*, *genius*, *pants*, *rubbish*, and others.

(d) Derivation by base modification

There are also cases where the extension of a base from one primary category to another is accompanied merely by a **phonological modification** of the base, not the addition of any affix. Usually the modification is a change in the vowel or final consonant, or a shift in the stress from one syllable to another. With minor exceptions, such modifications are not reflected in the spelling. We illustrate here with noun–verb pairs, marking stressed syllables where relevant by putting them in SMALL CAPITALS, but without attempting to indicate which form is the primary one:

[34]		NOUN		VERB
i	a.	*Here's the <u>document</u> you need.*	b.	*We will <u>document</u> the process.*
ii	a.	*I want to stay in this <u>house</u>.*	b.	*We must <u>house</u> them locally.*
iii	a.	*The interview was sheer <u>tórment</u>.*	b.	*They want to <u>torménte</u> me.*

- In [i], the noun and verb usually differ in the vowel of the last syllable. The last syllable of the noun has a reduced vowel, so it sounds the same as the last syllable of *informant*. The verb, however, has a full vowel, so it rhymes with *bent*.
- In [ii], there is a difference in the final consonant: the noun rhymes with *mouse*, whereas the verb rhymes with *rouse*.
- In [iii], the noun has the stress on the first syllable, while the verb has it on the second.

(e) Back-formation

A kind of opposite of affixation is found in some cases where a new word without an affix comes into the language later than a word consisting of the same base with an affix attached. From a historical perspective, therefore, we get a derived word that is formed by SUBTRACTING an affix from a base rather than adding one. This book is mostly not concerned with facts about the history of English, but this is one historical process that is worth noting.

- The verb *self-destruct* is a clear example. It was formed from the noun *self-destruction* by dropping the *·ion* suffix. Notice that the verb related to *destruction* is *destroy*, not **destruct*. The verb *self-destruct* came into the language after the noun, not the other way round. This backwards derivation of a word by removing an affix from an already existing word is known as **back-formation**.
- The classic example is the verb *edit*, which is known to have arisen through back-formation from the noun *editor*: again, the noun preceded the verb historically.

There is nothing in the forms themselves that enables one to distinguish between affixation and back-formation: it's a matter of historical formation of words rather than of their structure. In fact, we have already listed two compound verbs that are back-formations: *baby-sit* and *sleepwalk* in [31iii]. These arose by back-formation from the nouns *baby-sitter* and *sleepwalking*. Structurally they are compounds, since they consist of two bases, but the verbs did not arise historically by compounding.

(f) Clipping

Clipping is another minor process of word formation that removes part of a base (sometimes with a change in spelling for the part that remains), as in these examples:

[35]	FULL FORM	*delicatessen*	*microphone*	*helicopter*	*telephone*	*influenza*
	CLIPPING	*deli*	*mike*	*copter*	*phone*	*flu*

Deli and *mike* lose the last part of the original word; *copter* and *phone* lose the first part; and in *flu* both the beginning and the ending of the original base are lost, leaving a middle syllable (which in the original base is not even stressed).

Clipping should not be confused with back-formation. The two key differences are these:

- The clipped form does not differ in meaning from the original: it is merely a **variant** with the same meaning, usually a more informal one (though the degree of perceived informality changes over time: today *phone* is no more informal than *telephone*).
- What is removed in a clipping is typically not any kind of morphological element: **catessen*, **rophone*, and **enza*, for example, are certainly not bases or affixes or anything else (though *tele·* happens to be a bound base meaning "far away", seen also in *telepathy* and *telescope*).

(g) Blending

The process of **blending** is comparable to compounding, except that part of one (or both) of the source bases is dropped at the boundary between them. Examples are given in [36]:

[36]	SOURCE	*breath analyser*	*parachute troops*	*chocolate alcoholic*	*motor hotel*
	BLEND	*breathalyser*	*paratroops*	*chocoholic*	*motel*

In *breathalyser* the beginning of the second base (the *an* of *analyser*) is lost; in *paratroops* the end of the first base (the *chute* of *parachute*) is lost. In the others, both bases lose parts; in *motel* the *t* is common to both bases.

(h) Initialism

The final word-formation process we consider creates bases from the initial **letters** of a sequence of words (or, in a few cases, of parts of words). We call this process **initialism**. There are two subtypes (though people often confuse them): an **acronym** is formed by initialism in a way that picks initials that spell out a pronounceable word. An **abbreviation** is pronounced simply by uttering the names of the letters.

[37]				
	i	ACRONYM	*NATO* (*North Atlantic Treaty Organization)*	
			AIDS (*acquired immune deficiency syndrome*)	
	ii	ABBREVIATION	*CIO* (*Chief Information Officer*)	*UN* (*United Nations*)
			DNA (*deoxyribonucleic acid*)	*TV* (*television*)

- In acronyms the word spelled out by the letters is pronounced the way we would expect a word with that spelling to be pronounced: *NATO* rhymes with *Plato*, *AIDS* rhymes with *maids*.
- In abbreviations we just string together the letter names: *CIO* sounds like *sea eye owe*. (It would be possible to pronounce *CIO* as 'sigh-oh' or *UN* as 'un', but people don't. This means that when you first see an initialism in print, you may not always know whether it is an acronym or an abbreviation, and you might guess wrong. There is no way to tell until you hear someone say it.)

Initialisms are usually written with upper-case letters. The most common exceptions to this are abbreviations of Latin phrases: *e.g.* (*exempli gratia* "for example"), *i.e.* (*id est* "that is"), and others (dictionaries commonly contain lists of these).

Some words written in lower case originated as acronyms but are now not easily recognizable as such. They include *scuba* (*Self-Contained Underwater Breathing Apparatus*) and *radar* (*RAdio Detecting And Ranging*). These have been in the language for decades, and many speakers will be unaware of their origin as acronyms.

8.3 Productivity and lexicalization

A word-formation operation is said to be **productive** if it's still available for the creation of new words on the fly by speakers, and **non-productive** if it is not.

Affixation by means of such suffixes as *·able*, *·ness*, *·er*, or such prefixes as *un·* and *pre·* is productive. You can put these affixes on words that have only entered the language recently: *emailable*, *nerdiness*, *rapper*, *ungoogled*, *prexeroxed*.

But there are other affixes which are no longer productive in Present-day English, like the ones underlined in *bond·age*, *duck·ling*, *drunk·ard*, *en·able*, *inform·ant*, or *young·ster*.

Productive affixes have different **degrees of productivity** – they have differing ranges of bases they can attach to. Among suffixes forming nouns, for example, *·ness* is highly productive, *·ity* somewhat less so, while *·dom* is of very low productivity (though it is still a little bit productive, as evident from such a recent coinage as *yuppiedom*).

At the highest degree of productivity, it is unnecessary, and indeed not feasible, to list in the dictionary all the words formed by the process in question. For example, *·like* can be added very freely indeed to nouns to form such compound adjectives as *catlike*, *doglike*, *mouselike*, *wolflike*, *apple-like*, *banana-like*, *pear-like*, and it would be a mistake to try to include all words of this kind in the dictionary.

Words which couldn't be formed with their present meaning by means of operations still productive in the grammar today are said to be **lexicalized**: they absolutely have to be included in a dictionary. These include words formed through processes in the past that have not given rise to productive operations in the language as it is now, such as *drunkard*, *bondage*, and so on, because merely from having seen *drunk* and *bond* and *·ard* and *·age* we can't figure out what words can be made from them (**bondard*, maybe, or **drunkage*?) – or what the words would mean. However, lexicalization covers other cases too, notably:

- Words like *durable*, *knowledgeable*, *perishable*: although *·able* suffixation is a highly productive operation, what's productive is its use on transitive verb bases, not bound bases (*dur·*), nouns (*knowledge*), or intransitive verbs (*perish*).
- Some words have meanings not predictable from the combination of their component parts. The meanings have to be specified individually. For example, the salient meaning of *considerable* doesn't match that of numerous words like *achievable* ("can be achieved") or *climbable* ("can be climbed"): it doesn't mean "can be considered", but "large, significant, or notable". Similarly, the compound *loudmouth*, as we noted in §7.2, doesn't mean "mouth which is loud", but "person

> who talks a lot, typically one who gives offence". And *gentleman* doesn't mean "gentle man", but "man of chivalrous manner and good breeding or high social position" – or it may be just a courteous variant of *man*.

Information about words of this sort must appear in the dictionary, because no matter how well you are acquainted with the general principles of lexical morphology, you can't figure out what you need to know about them.

In general, the grammar of a language can only cover the rule-governed aspects of its ways of structuring words or sentences. That's what we've been concentrating on throughout the whole of this book. But to complete the description of a language, we would need similarly exact accounts of the pronunciation of sentences and the meanings of sentences, and finally a dictionary, to list the words in current use with their unpredictable phonological, morphological, syntactic, and semantic properties.

Exercises

1. Explain why *·ing* is an **inflectional** suffix in [i] but a **lexical** one in [ii]:

   i *They're building more town houses at the end of the street.*
   ii *We left the building by the back exit.*

2. Explain carefully why the suffix *·ish* (as in *greenish*, *sweetish*, *newish*) is NOT one of the inflectional suffixes of English.

3. Explain why there is **consonant doubling** in the first member of the following pairs, but not in the second:

   i *hopping* *hoping*
   ii *referring* *mothering*
   iii *stemming* *steaming*
   iv *quizzing* *boxing*
   v *starred* *stars*

4. For each of the following verbs say whether the final *e* is deleted or retained when the **gerund-participle** is formed. Relate your answers to the rule of *e* deletion discussed in §3.2.

   i ***age***
   ii ***be***
   iii ***centre*** [BrE]
   iv ***impinge***
   v ***implore***
   vi ***judge***
   vii ***plane***
   viii ***sortie***
   ix ***tinge***
   x ***wage***

5. For each of the following lexical bases give the **inflectional form** specified below, and show how the treatment of the final *y* follows the rule of final *y* replacement.

   i ***dry*** gerund-participle
   ii ***embody*** preterite
   iii ***guy*** plural
   iv ***silly*** superlative
   v ***try*** [V] 3rd sing present

6. Rewrite these examples with all noun phrases changed to their **plural** counterparts and all present-tense verbs changed to the correct **preterite** form.

   i *The other student sings in a rock band.*
   ii *The TV series made from that novel is as good as any film ever made.*
   iii *The bison roams the prairie and the wolf preys on the deer in the forest.*
   iv *The man drives the Mercedes into a garage and hopes no thief has a key to the building.*
   v *The chief focus of this task is investigating the larva and developing a criterion for distinguishing its response to an environmental stimulus from any similar phenomenon at a later stage.*

7. The following irregular verbs have **shape-sharing** between the **preterite** and the **past participle**. Assign them to one or other of Types A–H in [16] according to the relation between these forms and the lexical base, and note those which also have regular variants of the preterite and past participle.

   i ***bind***
   ii ***burst***
   iii ***dig***
   iv ***dream***
   v ***fight***
   vi ***hold***
   vii ***kneel***
   viii ***leave***
   ix ***lend***
   x ***meet***

8. As in the previous exercise, these are irregular verbs with shape-sharing between the preterite and the past participle. Assign to a type, noting those which also have regular variants of the preterite and past participle: [i] ***slide***; [ii] ***spell***; [iii] ***spread***; [iv] ***strike***; [v] ***weep.***

9. The following irregular verbs have distinct **preterite** and **past participle** forms. Assign them to Types A–H in [18] according to the way these forms are related to the lexical base.

   i ***fall***
   ii ***give***
   iii ***shrink***
   iv ***sow***
   v ***swell***
   vi ***get*** [AmE]
   vii ***grow***
   viii ***slay***

ix ***strive***
x ***wake***

10. For each word in the sentence *Our children said those earlier stories had been her worst*, say what **inflectional form** it is, and what **lexeme** it belongs to:

i *our*
ii *children*
iii *said*
iv *those*
v *earlier*
vi *stories*
vii *had*
viii *been*
ix *her*
x *worst*

(For example, if we had given the sentence *Pigs will fly*, then for the word *pigs* you would say that it is the plural form of the lexeme ***pig***. For inflection of personal pronouns, recall the analysis given in §5.8.)

11. Show for some plural-only words other than *scissors* that the **lexical base** sometimes occurs in compounds.

12. Discuss the choice between the *·es* and *·s* alternants of the plural suffix with the following nouns, after gathering evidence about how they are actually spelled in real texts:

i ***cameo***
ii ***echo***
iii ***eunuch***
iv ***garage***
v ***innuendo***
vi ***lunch***
vii ***mango***
viii ***patio***
ix ***photo***
x ***piano***

13. Nouns with lexical bases ending in *f* or *fe* either have (a) OBLIGATORY modification of the base in plural formation; (b) OPTIONAL modification; or (c) NO modification. Give plurals of the following nouns, grouping them into these three types:

i ***elf***
ii ***handkerchief***

iii ***life***
iv ***oaf***
v ***self***
vi ***sheaf***
vii ***shelf***
viii ***spoof***
ix ***waif***
x ***wharf***

14. Give plurals of the following nouns, grouping them into three types: (a) those with only **foreign** plurals; (b) those that have foreign and regular plurals as **variants**; or (c) those with only **regular** plurals.

i ***alumnus***
ii ***amoeba***
iii ***appendix***
iv ***crucifix***
v ***desideratum***
vi ***foetus***
vii ***mausoleum***
viii ***millennium***
ix ***phobia***
x ***radius***

15. Give **genitives** of the following NPs, grouping them into three types: (a) those with only **bare genitives**; (b) those that have ***'s*** and **bare genitives** as variants; or (c) those with only *'s* **genitives**: [i] *the children*; [ii] *Jones*; [iii] *these kids*; [iv] *that species*; [v] *Xerxes*.

16. Construct examples (grammatical and ungrammatical as the case may be) to show clearly that the adjective ***kindly*** can inflect for grade despite being formed with the *·ly* suffix, but the adverb *kindly* cannot be inflected for grade at all.

17. List all the **bases** that occur in the following words, bearing in mind that one base can be contained within another.

i *clothes-drier*
ii *handwriting*
iii *disinterestedness*
iv *self-righteously*
v *taxpayer-funded*
vi *injustice*
vii *upbringing*
viii *babysitting*
ix *unavoidability*

x *underachiever*

18. Discuss the form of the underlined **lexical bases** in the following examples: identify the **morphological operations** and the bases and affixes involved in their formation.

i *I got it from a CNN newscast.*
ii *It was formed by a process of adjectivalisation.*
iii *She works for UNESCO.*
iv *We didn't have lead-free fuel in those days.*
v *They are involved in some illegal operation.*
vi *They have decided to euthanize the whale.*
vii *Calm down.*
viii *He's always mouthing off about his boss.*
ix *I'll meet you in the lab.*
x *I was working as a window-cleaner.*