Chapter 4

Morphology: Word Structure

Francis Katamba

4.1 Morphology: the study of word structure

Morphology is the component of grammar that is devoted to the study of the internal structure of words. It is concerned not only with speakers' knowledge of the structure of existing, well-established words (e.g., *shopkeeper* is analysa-ble as *shop-keep-er*) but also with the rules used to form or interpret new words. Look at the words below. I suspect some of them will be new to you; they are relatively recent additions to the English lexicon. Explain how you might work out their meanings without resorting to a dictionary:

[4.1] earwitness f.i.n.e. McJob
televangelist wordrobe disorient express

If you're unsure about some of these words, help is at hand in Table 4.1.

You know many things about word-formation, including the fact that words have internal structure. New compound words can be formed by combining two existing full words, for example, *ear-witness* (the analogy with *eyewitness* is

obvious; and by combining chunks of words you can create a new blended word like *televangelist* (from *television-evangelist*). You know also that a new word can be formed by combining the first letters of words in a phrase as in *f.i.n.e.* or *imho*. In addition, you bring to the task of word structure analysis considerable shared cultural knowledge. For instance, you have views about the classiness (or otherwise) of *McDonald's* that inform your interpretation of the likely meaning of *McJob*. The intended wordplay on *wardrobe* in *wordrobe* will not escape you; in the compound *disorient express* you see the allusion to the *Orient Express*, a train eponymous with luxury long-distance travel (with a touch of intrigue). Much of this knowledge is subconscious. The study of morphology aims to make it explicit.

Morphology is the study of word structure. The problem is that although everyone knows what a word is, it is very difficult to give a simple definition of the word and to provide reliable criteria for recognizing words. A standard definition of a word is this: 'a word is a minimum free form' (Bloomfield, 1935: 178). This means that a word can occur on its own as an utterance. For instance, you could say, *Boys*! (with a sigh in exasperation, on seeing boys behaving badly). In fact, the word not only is the smallest form that can occur by itself as an utterance, it is also the smallest unit that can be manipulated by syntax. Consider the sentences below:

Table 4.1 Some relatively new words in English

EARWITNESS, n. An individual who hears an incident occur, especially one who later gives a report on what he or she heard [blend of ear and eyewitness].

Context and source: 'I didn't see what happened, I just heard it. I guess I was an earwitness.' (Conversation)

IMHO, n. acronym (In My Humble Opinion)

Context: Acronym used on the internet

F.I.N.E., adj. acronym (F*cked up, Insecure, Neurotic, and Emotional). Usually derogatory; indicated by tone of voice.

Context: 'We all know he's F.I.N.E, just like those freaks he hangs out with.'

MCJOB, n. A job in a service-related field with low pay, low prestige and little opportunity for advancement [analogous word formed from prefix Mc– and job]. Context and source: '… a message that I suppose irked Dag, who was bored and cranky after eight hours of working his McJob'. (Generation X, p. 5)

TELEVANGELIST, n. An evangelist who conducts regular religious services on television [blend of television and evangelist]

Context and source: 'Ole Anthony and his merry band take on the televangelists.' (US News, 8 December 1997)

WORDROBE, n. a person's vocabulary [blend of *word* and *wardrobe*].

Context and source: 'He has an extensive wordrobe.' (a web page on the internet)

DISORIENT EXPRESS, n. A state of confusion [novel formation].

Context and source: 'I felt like I was on the Disorient Express for good this time.' (*Newsweek*, 14 November 1996)

(Kemmer, 2004)

[4.2] a. The neighbour's dogs disturbed burglars.

b. Burglars were disturbed by the neighbour's dogs.

c. *s burglar weres disturb the ed by the neighbour's

d. *sdog the neighbour's disturbed s burglar.

e. The neighbour's *dog sdisturb burglar ed.

f. *Dog ed the s neighbour's disturb the s burglar.

Let us look more closely at some of these examples. We can say that *boys* and *disturbed* are words because they belong to the class of elements that are the smallest units that can be manipulated by syntactic rules. For instance, when passive sentences are formed, whole words (or phrases) are shunted around as you can see if you compare the active sentence in [4.2a] with its passive counterpart in [4.2b]. By contrast, the forms –s and –ed are not words. Neither of them can occur on its own, or be independently

relocated in a new position in the sentence by a syntactic rule. Hence the ungrammaticality of the examples in [4.2c]–[4.2f] respectively. Of course, phrases like *the burglar* and *the neighbour's dog* are free forms capable of occurring in isolation, or being moved from one position to another in a sentence. But they are not words since they are not the smallest units that syntax manipulates. They can be broken down into smaller free forms, namely the words *the*, *neighbour's* and *boys*.

ADVANCES BOX 2.1

ADVANCES BOX 4.1

Neologisms – *are you a shlumpadinka?*

Neologisms appear every day. Few of them become established words that the speech community broadly recognizes and which get listed in august dictionaries like *Merriam-Webster* or the *Oxford English Dictionary*. *Shlumpadinka* is one that seems to have made it. Slang words, especially those used by celebrities who are the high priests of popular culture, probably have a better chance of catching on than words like those listed in [4.1] that are created by anonymous speakers. Notice how this one introduces into English from Yiddish a consonant cluster in the *Shl* (/ʃl/) syllable onset that is not already found in the language (see also *Shm* – reduplication on page 53).

> shlumpadinka • \shlum-puh-DINK-uh\ • noun : a woman who dresses like she has completely given up on herself and it shows : a dowdy and unstylish woman

Example Sentence:

There you are running out to get the paper looking like a shlumpadinka. – Oprah Winfrey (27 April 2007)

Did you know?

'I have to practice not looking like a *shlumpadinka* on the air', said Oprah Winfrey on her eponymous show, broadcast 15 April, 1997. Oprah has occasionally separated *shlumpa* and *dinka*, the constituent parts of the word, for emphasis: 'You are watching right now in your sweats ... the same sweats you had on yesterday and the day before ... you are a *shlumpa* and a *dinka* and you know it!' Oprah has also used the word *shlumpadink* to refer to a masculine subject, although this form is somewhat less frequently heard. Oprah has often used *shlumpadinka*

attributively to modify another noun, as in 'It's my *shlumpadinka* shoes!' or 'You're watching me right now in your *shlumpadinka* pajamas.'

This is not the problem completely solved yet. We still need to clarify what we mean by **word**. The term 'word' is used in a variety of senses. These are summarized in Table 4.2.

4.2 Simple and complex words

Is a word the smallest unit of morphological structure? The answer has to be no. Take the word *boys*: it contains two meaningful structural units, *boy* and *–s* (plural). The minimal unit which has a meaning (e.g., *boy*) or a grammatical function (e.g., *–s* (plural)) is called a **morpheme**. Identify the morphemes found in each of the following words:

[4.3] dog childish undeservedly rewriting

Table 4.2 Distinguishing between the senses of the term 'word'

What is a word?
Word-form A word-form is a shape that represents a word in speech or writing. In the written language we call word-forms **orthographic words**. Thus, the last sentence has nine orthographic words. Normally, an orthographic word is preceded and followed by a space (or a punctuation mark). I say 'normally' because sometimes a compound word is

written with white space surrounding each of its constituent words (see *disorient express* in Table 4.1). Conversely, internet web-speak has produced a convention of writing web addresses 'as all one word' as in *www.alloneword.org/* when clearly what we have are separate words.

The word-form in the spoken language is called the **phonological word**. It is subject to various phonological constraints. For instance, as we saw in Chapter 3 (Section 3.2), all content words must have a syllable that bears primary stress. Furthermore, in a compound word one word has a syllable that outranks the rest in prominence.

Lexeme

The term 'lexeme' (or **lexical item**) is used when by 'word' we mean **vocabulary item**, i.e., an item requiring a single dictionary entry. Thus we can say that the word-forms *speak*, *speaking*, *spoke* and *spoken* are all different manifestations of the lexeme *speak*.

Grammatical word

Words viewed as syntactic units are called **grammatical words**. In this case we consider a lexeme together with the morphological and syntactic properties (morpho-syntactic properties for short) associated with it. For example *speaking*, *spoke* and *spoken* are characterized as grammatical words in this way:

- *speaking* is the **progressive** form of the verb *speak*
- *spoke* is the **past tense** form of the verb *speak*
- *spoken* is the **past participle** form of the verb *speak* (the past participle is the form of a verb that appears after the auxiliary verb *have*, as in *I have spoken*, *she has spoken*, *we had spoken*).

I expect that your solution is like the one in [4.4].

[4.4]

Words	Constituent
dog	dog
childish	child-ish
undeservedly	un-deserv-ed-ly
rewriting	re-writ-ing

We can distinguish between simple words which contain just one morpheme (e.g., *dog*) and complex words containing more than one morpheme, for example, *child-ish*, *un-deserv-ed-ly* and *re-writ-ing*. Observe that they can be classified as **free morphemes** or **bound morphemes**. A free morpheme can occur in isolation (as a word), for example, *dog*, *write*, *deserve* and *child*. But a bound morpheme cannot occur in isolation, for example, in [4.4] the forms *–ish*, *un–*, *–ed*, *–ly*, *re–*, *–ing* are bound morphemes.

Furthermore, any form that is used to represent a morpheme is called a **morph**. Thus the word *child-ish* has two morphs and *re-writ-ing* has three. Many mor-phemes are represented by more than one actual form. In other words, they have a number of variants that realize them in different contexts. Such vari-ants are called **allomorphs**. (The parallelism between morph, allomorph and morpheme on the one hand, and phone, allophone and phoneme on the other (Chapter 3, Section 3.1) should be obvious.) Allomorphs are distinct with regard to form. But they have the same grammatical or semantic function. Allomorphs are always in complementary distribution, just like allophones. Each allomorph is restricted to appearing in its allotted contexts. For instance, the indefinite article in English has two allomorphs *a* ~ *an*. (The symbol ~ means 'alternates

with' or plainly 'or'.) These two forms have exactly the same meaning. But they occur in different contexts (i.e., in complementary distribution):

[4.5] a. *a* is used if the next word starts with a consonant, e.g., *a leg, a mother, a tomato*
b. *an* is used if the next word starts with a vowel, e.g., *an ear, an aunt, an egg*

4.3 Word structure: a closer look

In this section, we shall consider the various types of morpheme, on the basis of their location and role in the word.

Root, base and affix

A **base** is a unit to which elements can be added in word-formation. For instance, starting off with the base *write*, we can add the bound morpheme *–ing* to form *writing*. Then *writing* can be used as a base to which we attach *re–* to form *rewriting*. An **affix** is a bound morpheme that must be attached to a base. If it precedes the base it is called a **prefix** and if it follows the base it is called a **suffix**. In [4.4] *re–* and *un–* are prefixes while *–ish*, *–ed*, *–ly* and *–ing* are suffixes. The base of a word may be a **root**. The **root** is the rump of a word that remains when all the affixes have been stripped away. *Write* is both a root and a base. But a base need not be a bare root. In many cases the base contains a root and one or more affixes; for example, we can take a root like *write* as the base, and add to it the prefix *re–* to obtain the word *rewrite*. This word, already containing a root and a prefix, can be used again as a base to which we add the suffix *–ing* to obtain the word *rewriting*. As mentioned above, at the start of this chapter, we can also form a compound word by combining two bases that are

words in their own right, for example, *ear* + *witness* gives *earwitness*.

Another major dichotomy is between **lexical morphemes** (also known as 'content words') and **functional morphemes** (also called 'function words'). Lexical morphemes are nouns, adjectives, verbs and adverbs (*NAVA* words in Chapter 5, Section 5.2). Functional morphemes, so called because they mostly signal syntactic relationships, include prepositions, pronouns and determiners (see Chapter 5, Section 5.7). An obvious difference that has important consequences for morphology is that lexical morphemes belong to an **open class** which can expand. But functional morphemes belong to an essentially **closed class** or set that admits no new members. New nouns, verbs, adjectives (and to a lesser degree adverbs) are created all the time. However, new prepositions, pronouns, determiners and the like are very rarely created. It follows that the branch of morphology that examines the creation of new vocabulary items is primarily concerned with lexical morphemes (see Chapter 14).

4.4 Inflection versus derivation

This leads us to the two broad classes of word-formation processes: **inflection** and **derivation**. In general terms inflection is syntactically motivated word- formation. For a word to appear in certain syntactic contexts, the syntax requires that it must have certain morphological marking, otherwise the sentence is ungrammatical. For instance, in Standard English if you have a third-person singular subject of a present tense verb, you must suffix *–s* to it, or else your sentence will be ungrammatical:

[4.6] a. She sleeps
Othman sleeps
It sleeps

b. *We sleeps
*They sleeps
*You sleeps

Unlike inflection, **derivation** is not motivated by the syntax. Rather, its role is to create new lexical items. Derivation involves one or both of these two things: creating a new lexical item with a different meaning from that of the original word or changing the syntactic class of the input lexical item, as seen in [4.7]:

[4.7] a. **Derivation changing meaning**

Input	Derived word
possible	impossible
tell	retell
do	undo

b. **Derivation changing syntactic category**

faith (noun)	faithful (adjective)
fierce (adjective)	fiercely (adverb)
sing (verb)	singer (noun)

In this section we will focus on the notion of inflection and the general morphological properties of inflectional morphemes, and we return to derivation in Section 4.6.

English has a very small number of inflectional morphemes, and they are all suffixes. Inflectional suffixes form a closed set. The language no longer adds to its inventory of inflectional endings. But historically English used to have considerably more complex inflectional morphology (see Chapter 16). Inflectional suffixes are listed in Table 4.3. The next chapter provides an opportunity to see how these inflectional suffixes are used in grammar.

The fact that inflection is syntax driven has already been highlighted. Many inflectional processes involve **agreement**. This refers to cases where the head of a grammatical unit requires elements in construction with it to take on certain morpho-syntactic properties of the head. In [4.8], for example, the morpho-syntactic property number that is associated with the head of the noun phrase (NP) (see Chapter 6, Section 6.5) dictates the choice of

inflection of the demonstrative. The demonstrative must agree in number with the head noun:

[4.8]

Singular	Plural
this boy (*these boy)	those boys (*this boys)
that boy (*those boy)	those boys (*that boys)

Another example of inflection is subject–verb number agreement in Standard English as illustrated in [4.6] above.

Finally, inflectional properties may be inherent. A morpheme is associated with those properties always regardless of context. For instance some nouns have the property [+count] and can be marked plural while others are not countable and lack a plural form, for example, *hammer* has the plural form *hammers* but *equipment* has no plural (**equipments*) (see Chapter 5, Section 5.5.1).

Table 4.3 Inflectional suffixes

Inflectional suffixes	*Orthographic form*		*Examples*	
PLURAL NOUNS				
Native regular plurals	–s		*toy*	*toys*
Native irregular plurals	/f/ voicing		*leaf*	*leaves*
	internal vowel		*mouse*	*mice*
	–en		*ox*	*oxen*
	zero	*sheep*	*sheep*	
	Sing.	**Pl.**		
Borrowed (all) irregular plurals	–us	–i	*cactus*	*cacti*
	–a	–ae	*amoeba*	*amoebae*
	–ex/-ix	–ices	*matrix*	*matrices*
	–on	–a	*automaton*	*automata*
	–eau	–eaux	*bureau*	*bureaux*
	zero	zero	*chassis*	*chassis*
	–o	–i	*concerto*	*concerti*
	zero	–im	*seraph*	*seraphim*
GENITIVE NOUNS	–s		*student's*	
VERBS				
third-person, present, singular	–s		*cooks*	
Progressive	–ing		*cooking*	
Past tense: Regular	–ed		*cooked*	
Past tense: Irregular	zero		(*let*→) *let*	
	(various) usually with internal vowel change		(*sing* →) *sang*	
			(*take* →) *took*, etc.	
Past participle	–ed, –en		*cooked, driven*	
ADJECTIVES				
Comparative degree	–er		*longer*	
Superlative degree	–est		*longest*	

In forming a word, a lexical base to which inflectional morphemes are attached (e.g., *sleep* in *sleep-s*) is called a **stem**. In a complex word which con- tains a sequence of suffixes, the derivational suffix is normally part of the stem. So, derivational suffixes are nearer the root and inflectional ones are on the outside, as in *sing-er-s* where *–er*, the suffix deriving an agentive noun (one who does X), is next to the root and the plural suffix is at the edge of the word.

Another difference, to be discussed in more detail below, is that inflectional morphology is for the most part regular, but derivational morphology tends to be sporadic. Almost all inflectional stems that are eligible to take an inflectional suffix do so. But rarely do a majority of bases eligible for a derivational process successfully undergo it. For instance,

virtually all verbs have an *–ing* (e.g., *writing*). But in derivational morphology it is the norm for rules to apply patchily, for example, you can *whiten* or *darken* things but you cannot **greenen* or **yellowen* them. If a method of word-formation is actively used to form new words it is said to be 'productive'. So, we can say that, by and large, inflectional processes tend to be more productive than derivational ones. True, some of the irregular variants of the inflectional morphemes in Table 4.3 are virtually unproductive (e.g., zero plural). But all the regular variants are very productive. We can suffix *–ing*, *–ed*, *–s*, etc. to new forms. If we tried to do the same with a derivational suffix, for example, *–fy*, we would produce acceptable words like *purify*, *glorify*, *terrify*, *gentrify*; but we must avoid getting carried away. Otherwise we would create words like **diligen-tify*, **accountantify* and **joyify*, which are not allowed. The line between what is allowed and what isn't is fuzzy. What do you make of *complexify*?

4.5 Morphological processes

Let us now survey some common morphological processes used in word-building.

Conversion

When **conversion** (also called 'zero-derivation') is used, a new word is formed by assigning an existing word a new syntactic category without changing its form in any way. You can tell which word class the derived word belongs to by looking at the syntactic context in which it appears.

[4.9] a. Put it in the bin$_{[noun]}$
Take a fast$_{[adjective]}$ train
He has short legs$_{[noun]}$

b. Bin$_{[verb]}$ it!
The train goes very fast$_{[adverb]}$
We legged$_{[verb]}$ it

Conversion is a very productive method of deriving words, especially verbs from nouns and nouns from verbs.

Affixation

Predictably, **affixation** is word-building involving the use of affixes. This is extremely common. Affixes can be classified as **prefixes** and **suffixes** depend-ing on whether they precede or follow the root, as we have already seen in Section 4.3. Another type of affix is the **infix**. This is an affix inserted in the root itself. Thus in the Alabama language of east Texas the second-person singular is indicated by the infix *–chi–* as in *ho-chi-fna* 'you (singular) smell' (from *hofna* 'to smell') (Stump, 2001: 131). In English, infixation is very marginal. Expletive infixation is used quite routinely for expressive purposes as in the examples in [4.10] from McCarthy (1982) but it has no grammatical role.

[4.10] Mononga-'fuckin-hela

[Monongahela is a river in Pittsburgh, Pennsylvania]

Ala-fuckin-'bama

fan-fuckin-'tastic

Dub-fuckin-'rovnik

in-fucking-'stantiate

Other expletives such as *bloody* and *friggin* can also be infixed.

If one fancies a bit of colourful language, how does one decide when and where to deploy an expletive, you may ask. According to McCarthy (1982), the expletive infixation phenomenon is subject to a phonological constraint that is best stated in terms of metrical foot structure. The word in

which the infix is inserted must be at least three syllables long and the infix must go inside the word after the syllable preceding a trochaic metrical foot (i.e., a foot with a stressed syllable followed by an unstressed syllable). Hence the unacceptability of **Glas-fuckin-gow* (from disyllabic *Glasgow*) and **Jo-friggin-hannesburg* (from four-syllable *Johannesburg*). This is a good example of how phonological structure can play a key role in determining if and how a word-formation operation applies. It is a good example of how you cannot describe word-formation without taking into account phonological information.

Zero morph

In Table 4.3, under the list of allomorphs of the noun plural morpheme, I listed zero as one of the possibilities. This was to cover the case of the word *sheep* which, unlike other nouns, gets no overt marking of number even when its meaning is plural as in *Twenty sheep were stolen yesterday*. Clearly *twenty* indicates plurality. The auxiliary verb were, which is in the plural, indicates that English syntax recognizes the plurality of this noun, which is in the subject noun phrase and requires the verb agreeing with it to be plural. Likewise many monosyllabic verbs, for example, *let*, *cut*, *hit*, are not marked overtly in any way when they occur in a context where the syntax requires past tense inflection on the verb, for example, *Yesterday I cut it* (**Yesterday I cutted it*). For this reason zero is listed as one of the allomorphs of the third-person singular past tense marker. Recognizing zeros, as we have done, makes sense because we have made the theoretical assumption that morphemes realize or represent the abstract concepts that lie behind morphemes. It wouldn't make much sense if we assumed, rather, that morphemes are actually made up of phonemes. (See the discussion of exponence below.)

Internal change

Sometimes inflection is done by changing a vowel in the root. This is called **internal change** (or 'apophony'):

[4.11] a. **Nouns**

Singular	Plural
foot	feet
[fʊt]	/fiːt/
mouse	mice
[maʊs]	[maɪs]

b. **Verbs**

Present tense	Past tense	Past participle
ride	rode	ridden
[ɹaɪd]	[ɹəʊd]	[ɹɪdən]
sing	sang	sung
[sɪŋ]	[sæŋ]	[sʌŋ]

In some cases, apophony may be accompanied by affixation as in the case of *ridden*.

Exponence

The key point to note is that the relationship between morphs and morpho-syntactic features such as plural, past tense, present tense, etc. is one of realization (or representation) rather than composition. Morphs belonging to a given morpheme represent one or more morpho-syntactic features. Matthews, in a number of publications going back to the early 1970s, proposed characterizing this relationship using the notion of **exponence** (Matthews, 1991: 175). He distinguishes different types of exponence:

(i) **Simple exponence** Simple exponence occurs when one morph (form) realizes a single morpho-syntactic feature, for example, [s] in *sweets* realizes plural, and [t] (the actual pronunciation of *–ed*) in *parked* realizes past tense or past participle.

(ii) **Cumulative exponence** A single morph realizes more than one morpho-syntactic feature simultaneously. Take the example of [s] in *she thinks*. The [s] of the verb represents third person, present tense and singular.

Suppletion

Normally allomorphs of a morpheme are phonologically related. Thus, the regular past tense ending in English is realized as [t] after a verb whose last sound is voiceless, and as [d] after a verb whose last sound is voiced (cf. *parked* [pɑːkt], *missed* [mɪst], *watched* [wɔːtʃt] vs. *lived* [lɪvd], *ruled* [ɹuːld] and *spied* [spaɪd]). The sounds [t] and [d] are quite similar, both being alveolar stops.

Occasionally we find allomorphs of the same morpheme whose phonological shapes are unrelated. If a phonological relationship is totally non-existent, we speak of **total suppletion**. The words *good* and *better*, *go* and *went* are examples of total suppletion. The term 'partial suppletion' is used to describe situations where residual phonetic similarity between allomorphs can be detected, as in the case of the verb *seek ~ sought*; *bring ~ brought*, etc.

Syncretism

The term **syncretism** refers to a situation where morpho-syntactic categories that are represented by distinct forms elsewhere are mapped onto a single form in some contexts. This was obliquely noted in the discussion of exponence above. In many regular verbs, and some irregular ones, the morpho-syntactic properties past tense and past participle are mapped onto different forms as shown in [4.12a]:

[4.12] a. No syncretism:

Past	Past participle
rode	ridden
gave	given
sang	sung

b. Syncretism:

Past	Past participle
cooked	cooked
listened	listened
brought	brought

ADVANCES BOX 2.1

ADVANCES BOX 4.2

Inflection vs. derivation – the importance of function

A matter related to the discussion of syncretism is the overlap in the realization of distinct inflectional and derivational morphemes. In Table 4.3, we noted that *–ing* is an inflectional suffix marking the progressive form of the verb. The use of *–ing* to represent the progressive is illustrated in [4.13a]. Contrast that with its use in [4.13b] where it is used to mark the gerund. (A gerund is a nominal form that comes out of a verb as a result of a derivational operation.) *Taking cocaine* is a noun phrase headed by the gerund *taking*. We can use the standard syntactic test of substitution to confirm this. In syntax we can only replace like with like, without affecting grammaticality. We can replace *Taking cocaine* with a noun (noun phrase) like *Success*, so *Taking cocaine* is also a noun phrase.

[4.13]	a.	*He is taking cocaine.*	Inflection : *taking* progressive form of verb formed by suffixing *–ing*
	b.	*Taking cocaine makes users feel on top of the world, allegedly.*	Derivation: *taking* gerund formed from verb by suffixing *–ing*

The moral of the story is that to determine whether a suffix is inflectional or derivational we need to examine how it is used in the syntax. Shape alone is insufficient.

Without syncretism there is internal change and suffixation of *–en*, or both, to signal past participle as seen in [4.12a]. With syncretism, exactly the same form, i.e. *–ed* is suffixed, only the context can help distinguish between past participle and past tense.

Haplology

Avoidance of sequences of identical linguistic forms is a phenomenon found in many languages. When two identical or very similar syllables or sounds occur next to each other it's not unusual for one to be deleted. This is called **haplology**. It can happen internally within a word or root morpheme of at least three syllables if a weakly stressed syllable is next to an adjacent syllable that is virtually

identical, as in *probably* [pɹɒbəbli] → *probly* [pɹɒblij]. Likewise, when a singular noun or a plural noun ending in [–s] or [–z] is put in the genitive, haplology occurs. Since the genitive suffix is spelled <s> and is phonologically realized as [s] ~ [z] (or [ɪz]), it is too similar to the final sound of the base for comfort. So, it is elided in many people's pronunciation and is normally omitted in the written form of the language (cf. *Jones's house* → *Jones' house* , *Charles's agents* → *'Charles' agents* , etc.). The genitive in these situations is indicated merely by the presence of the apostrophe.

Table 4.4 *Noun–verb stress doublets*

Noun *Initial stress*			*Verb* *Final stress*		
'reject	'progress	'contract	re'ject	pro'gress	con'tract
'produce	'refuse	'digest	pro'duce	re'fuse	di'gest
'import	'insult	'escort	im'port	in'sult	es'cort

Stress placement

In a minority of cases, derivation is effected by changing stress placement. Nouns can be derived from verbs and verbs from nouns by the rule introduced above (Chapter 3, Section 3.4).

Reduplication

Reduplication is the creation of a new word by repetition of an existing word in its entirety, or in part. Repetition of the entire word is called **full reduplication** to contrast it with **partial reduplication** where only part of a word is repeated. Full reduplication is exemplified by words like *bang bang, bye-bye, night-night* and *go-go*. Partial reduplication has traditionally involved rhyming, as in *airy-*

fairy, hoity-toity, razzle-dazzle, nitty-gritty and *willy-nilly,* or ablaut as in *tip-top, shilly-shally, zigzag, pitter-patter*. The process affects a monosyllabic word. The vowel of the rhyme is changed, leaving the rest of the word intact. In addition, there is a highly productive 'shm-reduplication' pattern (see Advances Box 4.3 below).

ADVANCES BOX 4.3

Shm-reduplication

None of the types of reduplication described above is very productive. By contrast, shm-reduplication, a phenomenon that has attracted considerable interest among linguists of late, is quite productive. Shm-reduplication is not very new; it has been in the language for over half a century. According to Nevins and Vaux (2003), it came into English from Yiddish as a result of language contact. In Yiddish, many words beginning with *s(c)hm–* have negative connotations. The pejorative use of *Shm–* was extended to English by Jewish speakers in the north-east of the United States. Gradually it spread to other dialects.

Shm-reduplication derives its name from the fact that *shm–* is used as the onset of the first syllable of the reduplicated word. The meaning of the reduplicated word tends to have uncomplimentary connotations (e.g., *You call yourself a captain? Captain shmaptain, what leadership did you show? You just cried when we went a goal down*).

Shm-reduplication is subject to a variety of phonological constraints. They are enumerated as follows by Vaux and Nevin (2003):

(i) If the first syllable of a word contains a lone consonant, that consonant is replaced by *shm–*, e.g., *baby shmaby, plan shman* and *table shmable*.

(ii) If a word starts with a consonant cluster some speakers just replace the initial consonant of the cluster while retaining the rest, for example, *breakfast shmreakfast*, while others replace the entire onset cluster with *shm–*, e.g., *plan shman, gravity schmavity*.

(iii) If a word starts with a vowel, *shm–* is put right at the beginning of the reduplicated form, which thereby acquires an onset to its first syllable, for example, *opinion shmopinion*, *apple shmapple*, *optician shmoptician* (see Vaux and Nevins (2003) for further discussion).

Observe that [ʃm]– is not a permissible onset in the native phonology. But language contact between Yiddish and English has resulted in the introduction of that syllable onset into English. Also, observe that shm-reduplication illustrates another important general principle, namely the

fact that certain morphological rules are very closely bound to morphological operations. We will return to this in the section below.

4.6 Derivation

As we noted in Section 4.3, derivation creates a new lexical item with a new meaning or grammatical category. In this section we will examine more closely the derivational morphology of English. The processes used are of three types: **affixation**, **conversion** and **compounding**.

4.6.1 Affixation

The most important method of derivation is **affixation**. Table 4.5 lists examples of English derivational affixes. Many well-established affixes are productively used to create new words, for example, the word *retraditionalization* that I saw in an advertisement for a sociology conference at the university.

The set of derivational suffixes is large, and unlike the list of inflectional suffixes, it can have new members added to it (see Table 4.5). I will illustrate the open-endedness of the derivational affix inventory with the suffix *–lite*. The word *lite* (a play on 'light') entered the language as part of the compound word *Miller Lite*, the brand name of a beer launched on the US market in 1967. Soon *lite* was appropriated by many food and beverage manufacturers, who used it to indicate that their products were low in calories and hence good for you. Subsequently, *lite*, which had been an independent word, was converted into a suffix and attached to bases as *–lite*. With regard to meaning, the favourable connotations of the suffix *–lite* were eventually eclipsed by ironic usage. I recently heard *marriage-lite* being used to refer to a marriage regarded by a speaker as

a sham (see also *grief-lite*, *interview-lite*, etc.; Dent, 2004: 22).

4.6.2 Conversion

Conversion is a very widely used method of forming words in English. When conversion occurs, the syntactic context is the only indicator that word class has changed. Conversion of N → V, and to a lesser extent V → N, is very productive. But that is not to say that anything goes. For example, deriving a verb from the noun floor works, but attempting the same with ceiling doesn't. Table 4.6 gives examples of word class conversion.

Table 4.5 Derivational affixes

	Affix	*Change*	*Examples*
a.	**Suffix**		
	–able	V→Adj	*likeable, teachable*
	–acy	Adj→N	*supremacy, celibacy*
	–ant	V→N	*celebrant, applicant*
	–al	V→N	*betrothal, denial*
	–al	N→Adj	*universal, educational*
	–ee	V→N	*payee, employee*
	–er	V→N	*driver, teacher*
	–ful	N→Adj	*pitiful, dreadful*
	–ic	N→Adj	*microscopic, philanthropic*
	–ing	V→N	*surfing, ironing*
	–ing	V→Adj	*sailing, travelling*
	–ish	N→Adj	*childish, boorish*
	–ity	Adj→N	*timidity, vanity*
	–ment	V→N	*government, amusement*
	–less	N→Adj	*shameless, joyless*
	–ness	Adj→N	*boldness, friendliness*
	–ly	Adj→Adv	*usually, quietly*
b.	**Prefix**		
	a–	Adj→Adj	*amoral, apolitical*
	dis–	V→V	*disagree, dislike*
	ex–	N→N	*ex-student, ex-chairman*
	re–	V→V	*re-write, re-sell*
	un–	Adj→Adj	*uncooked, unclean*

Table 4.6 Conversion

V→N	N→V	Adj→V	Adj→N	Adj→Adv
walk	*chair*	*sour*	*weekly*	*slow*
throw	*floor*	*cool*	*comic*	*fast*
kick	*plant*	*green*	*stiff*	*dead (as in dead slow)*
dig	*bottle*	*wet*	*wet*	
kill	*text*	*open*		
cry	*red card*			
read	*google*			

Sometimes, we are unsure which way the derivation went historically, for example, is *plan* a noun derived from a verb or a verb derived from a noun?

4.6.3 Compounding

Compounds are complex words containing at least two bases that are themselves words. It has always been a highly productive process in English (see Chapter 14). Normally compounds are classified on the basis of the word class of their constituents and the class of the entire resulting word. Compounds always have a headword which assigns its syntactic properties to the entire word, and thanks to the **right-hand head rule** – the idea that the head is the right-hand-most element – it is normally governed by the right-hand-most word. Examples of English compounds are given in Table 4.7.

It is very common for words formed by affixation to be part of a compound. Normally, the right-hand head rule applies and the last word in the compound assigns its class to the entire word as you can see in the tree in Figure 4.1. Compounds can also include other compounds, some of which may have affixes. This results in words of considerable complexity such as *armchair sportsman* (see Figure 4.2).

Clearly, the right-hand head rule is very important in English word- formation. But it doesn't hold absolute sway. It is subverted in two major ways. First, as we have seen, English readily allows conversion. So a word's syntactic

make-up may not reflect its syntactic category as indicated by the morphemes it contains, for example, *up market* (as in *an upmarket pub*) is an adjective not a noun, though the last word of the compound is a noun. Second, phrasal verbs (e.g., *look for, look up to, look into*) are a very large class of compound words with a head on the left.

Having said that, we still have to recognize the fact that many compounds are **compositional**. Their meaning can be inferred from the meaning of the words they contain. The word on the left serves as a modifier of the headword that is on the right. So they needn't be listed as separate headwords in the dic-tionary; they don't need to be memorized. Such compounds are called **endocentric compounds**, for example:

Table 4.7 Examples of compound words

a. **Noun compounds**	
N + N	**A + N**
$[school]_N\ [boy]_N \rightarrow [schoolboy]_N$	$[wet]_A\ [lands]_N \rightarrow [wetlands]_N$
$[gun]_N\ [dog]_N \rightarrow [gun\ dog]_N$	$[strong]_A\ [man]_N \rightarrow [strongman]_N$
b. **Verb compounds**	
N + V	**Prep + V**
$[house]_N\ [train]v \rightarrow [house\ train]_v$	$[under]_{Prep}\ [sell]_V \rightarrow [undersell]_V$
$[speed]_N\ [date]_V \rightarrow [speed\ date]_v$	$[over]_{Prep}\ [reach]_V \rightarrow [overreach]_V$
c. **Adjective compounds**	
Adj + Adj	**N + A**
$[yellow]_{Adj}\ [green]_{Adj} \rightarrow [yellow\ green]_{Adj}$	$[razor]_N\ [sharp]_{Adj} \rightarrow [razor\text{-}sharp]_{Adj}$
$[red]_{Adj}\ [hot]_{Adj} \rightarrow [red\text{-}hot]_{Adj}$	$[war]_N\ [weary]_{Adj} \rightarrow [war\ weary]_{Adj}$

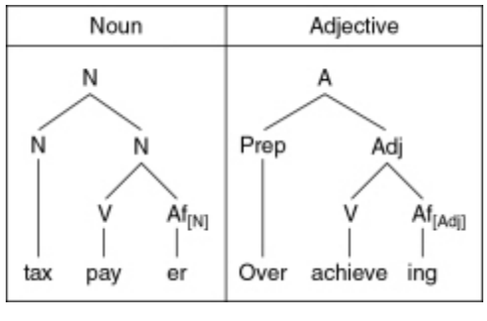

FIGURE 4.1 Compounds including affixed words

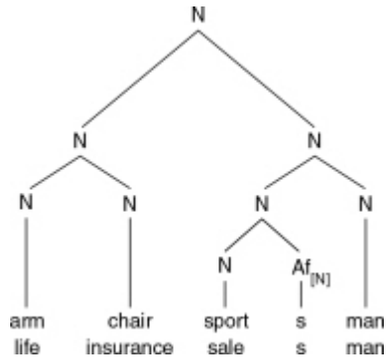

FIGURE 4.2 Compounds containing an affixed compound

[4.14] a. frying pan, dancing shoes b. can opener, candlestick maker

Endocentric compounds must be distinguished from compounds like those in [4.15] whose meaning is not compositional. Such words are called **exocentric compounds**. They must be listed in the lexicon; they must be memorized.

[4.15] green tax (N), greenhouse (N)

cold turkey (N), bag lady (N)

golden handcuffs (N), stonewall (V)

cold shoulder (N,V), carbon footprint (N)

Of course, there is the added complication that the same expression may be an exocentric or endocentric compound, depending on context. Take the adjective *knee-deep*. It can mean literally 'sunk right up to the knees', in which case it is compositional and hence endocentric; or it can be figurative, meaning 'deeply involved with', in which case it is noncompositional and hence exocentric.

Now, here is another conundrum. Compound words are made by combin-ing words. And so are syntactic phrases. So, how do we distinguish compound words from syntactic phrases? To clarify, the issue here is distinguishing

between, say, the compound *greenhouse* (the structure in which vegetables are typically grown) and the phrase *green house* (a house that is coloured green). One of the clues we use is stress. English has a special rule for assigning stress to compounds which doesn't apply to syntactic phrases. Compare the words in [4.16]:

[4.16]	**Compounds**	**Phrases**
	'blue,blood	,blue 'blood
	'green,house	,green 'house
	'New,house	,new 'house
	'Old,man	,old 'man

Various criteria for identifying compounds have been proposed. Some are phonological. For instance, in Adj + N compounds like those in 4.16, the primary stress of the compound is on a syllable in the first word; the main stress of the second word is subordinated to it. But in phrases the reverse is true. In the written language a word has white space before and after it. Often this is true of compounds. Either they are written without a space between them (e.g., *sunset*) or they are hyphenated (e.g., *double-sided*). But very many compounds are written as separate, unhyphenated words (e.g., *life insurance*, *Prime Minister*). (Revisit the discussion of word-forms in Table 4.2.) So, we need to rely on syntactic criteria in many cases. If two or more words are treated as a single unit for syntactic purposes, for example, if they can be moved, elided or substituted for as a single unit by the syntax, they constitute a compound. Another clue is inflection: normally, if a unit constitutes a word, any inflection it receives appears at its right edge (courtesy of the right-hand head rule). Thus, the plural of *mousetrap* is *mousetraps*

(*micetrap*); the plural of *bean feast* is *bean feasts* (**beans feast*, etc.).

ADVANCES BOX 4.4

Cranberry words

The impression that might have been given so far is that the difference between compounding and affixation is always clear-cut. Reality is more complex. The norm is to recognize affix morphemes as word-building elements that recur in many words, for example, *re–*, *–er*, *–ing*, *non–*. This principle serves us well in most cases. But from time to time the evidence is indeterminate. If a form appears in only one word, should we recognize it as a bound morpheme? Should we do so even if its meaning is totally obscure?

These questions arise when we consider the words in [4.17]:

[4.17]	a.	blackberry blueberry	b.	strawberry raspberry gooseberry	c.	cranberry huckleberry mulberry

The words in [4.17a] are clearly compounds; they are made of the words *black*, *blue* and *berry*. The words in [4.17b] also appear to be compounds, for the same reason. But there is a problem here. Present-day speakers of English for the most part do not comprehend the meaning and relevance of *straw*, *rasp* and *goose* in these words and view them as simple, mono-morpheme words. The problem is even more severe in [4.17c]. The bound morphs *cran–*, *mul–* and *huckle–* occur only in these forms in the entire language. Their specific meaning is elusive. Nor is it clear whether they are root morphemes or prefixes. If they are root morphemes, the words in [4.17c] are compounds. Otherwise, in [4.17c] what we have is a prefix followed by a root. Should a form that is encountered just once in the language and which has no clear meaning be recognized as a morpheme? Those with a cautious disposition would and do argue for treating such forms as simple words with a single morpheme.

The knotty issues raised by cranberry words are of wider relevance. There are many other situations where the zeal with which the linguist identifies morphemes has to be tempered. It is uncontroversial to say that *dislike* and *disagree* contain two morphemes each. But what about *disgruntled*, *disturb* and *dismay*? If *dis–* is identified as a morpheme, the forms *–gruntled* (from *–gruntle*?) and *–may* cut unconvincing figures as morphemes since neither of them recurs nor has an independent meaning. For analogous reasons, *highfalutin* is problematic. Segmenting it into *high* and *falutin* gives a compound with a dubious second element.

There are many words borrowed from Latin or Greek with a root whose meaning classical scholars can identify but which is hidden from the rest of us. For instance, the study of etymology shows that the root *–pol–* from a

Greek word meaning ‘city’ or ‘state’ is found in words like *police*, *politic*, *policy*, *monopoly*, *metropolis*, etc. Should *–pol–* be recognized as a root morpheme in all those words in a description of present-day English? I would suggest not. Since the task of the linguist is primarily to describe what current speakers of a language know about their language, there is no justification for identifying forms as morphemes if speakers in general are completely unaware of them and they have no relevance to their linguistic competence.

4.7 Further sources of English words

Derivation using affixation, conversion and compounding are the principal methods used to create vocabulary items. But there are also other methods. They are sketched out below.

Coinage

Word manufacture without recycling existing words and morphemes is called **coinage**. It is rare. It is mostly found in names of corporations and their commercial products; for example, according to the *Oxford English Dictionary*, *nylon* is an invented word. It was coined in 1938 by the DuPont pharmaceutical company by combining the fabricated stem *nyl–* with the pseudo-suffix *–on* found in other names of textile fibres (cf. *rayon*, *cotton*). Likewise, the New Yorkers Reuben and Rose Mattus fabricated the name *Häagen-Dazs* for their ice cream brand.

Eponyms

Eponyms are new words created by widening the meaning of a personal name to refer to a place, concept or product associated with that person, for example, *Kafkaesque* (from Franz Kafka); *lynching* (from Charles Lynch); *mesmerize*

(from Franz Mesmer); *pasteurization* (from Louis Pasteur); *Seattle* (from Chief Seattle), *Victoria* (in Australia, British Columbia, etc., from Queen Victoria). Business people are normally no shrinking violets and often their big egos are projected in the names of their companies and their products. This makes eponymy a highly productive word-formation process as new companies and products come into existence all the time. Thus, *Boeing*, the aircraft manufacturer, is named after its founder, William E. Boeing, whose name replaced the original name, Pacific Aero Products Co.; Michael Dell made sure that his computer company, and its products, were called *Dell*, and so on.

Backformation

Normally word-formation involves addition rather than subtraction. Affixes are added to a base, or two words are combined to form a word. The processes that we are going to be considering in this subsection, and in a subsection following it, all involve taking something away from the input. We will start with **backformation**, which arises from a reinterpretation of the structure of a word so that a chunk that is re-analysed as an affix is removed, leaving behind the assumed root. In the examples in [4.18], which are borrowed from Williams (1975), the schwa of the rhyme of the final syllable (spelled –er ~ –ar ~ –or) was re-analysed historically as the agentive suffix that forms nouns from verbs (as in sing > singer). It was removed yielding the words in [4.18b]:

		a. Original word	b. Backformation
[4.18]		beggar	beg
		peddlar	peddle
		hawker	hawk
		scavenger	scavenge
		editor	edit

In the more recent past and in present-day English, backformation has been used to derive *automate* from *automation*, *adulate* from *adulation*, *auto-destruct* from *auto-destruction*, *bulldoze* from *bulldozer*, *choreograph* from *choreography*, *babysit* from *babysitter*, *text*$_{(V)}$ from *text message*.

ADVANCES BOX 4.5

Backformation on the march in current journalistic writing (after Neal, 2006)

There are interesting developments in the use of backformation in current journalistic writing, extending the process in an innovative way, as the excerpt below explains:

> The backformations that usually catch my eye (or ear) are those in which a noun or adjective form of a verb (i.e., a gerund or participle) is compounded by putting a noun or adjective in front of it, and then the –ing or –ed is stripped off to yield a new verb. For example, in *performance-enhancing drugs*, we have a noun+adjective compound (*performance* plus *enhancing*), and the whole thing functions as an adjective, modifying *drugs*. But in Ruben Bolling's *Tom the Dancing Bug* comic strip of 18 Dec. 2004, one character tells another:
>
> > You've been performance-enhancing!
>
> Sure, it still sounds the same, but now *performance-enhancing* is the progressive form of a verb, not an adjective modifying something else. In the future, look out for finite verb forms, such as *He performance-enhanced illegally*, *None of our players performance-enhance*, or *Everyone performance-enhances!*
> Other examples of this kind of backformation where a noun glued to the beginning of the verb takes the place of a direct object after the verb:
>
> - Adam's therapist mentioned that he 'emotion-shared' at some point.
> - 'The decision puts Limbaugh back near square one and is likely to reinvigorate the criminal investigation into whether he "doctor-shopped"' ('Justices won't hear Limbaugh appeal', *Palm Beach Post*, 29 April 2005).
> - From a Bionicle comic in a Lego magazine: 'We came, we scouted, we all-conquered.'

Blending

Chunks of words may be **blended** to form new words. Less commonly, and mainly in the realm of IT, it is the initial chunks of two words that are combined, for example, *modulator* **+** *demodulator* gives *modem*; Wireless Fidelity gives *Wi Fi*. More commonly, the initial chunk of the first word is combined with the final part of the second word, for example, *brunch* (*breakfast lunch*), *insania* (*insanity mania*), *chugger* (*charity mugger*), etc. *Adidas* was formed from a blend of the name of the company founder Adolf (*Adi*) + *Dassler*. Thus it exemplifies the use of both eponymy and blending.

Clipping

Shortening long words by dropping a part is called **clipping**. Some clipped forms like *fab* (from *fabulous*) and *brill* (from *brilliant*) are slang but others like *bus* (from *omnibus*) and *gym* from (*gymnasium*) are very much part of the standard language. We can distinguish between three types of clipping with regard to structure. First, there is 'fore clipping', that is, deleting the first part and keeping the final part, for example, *varsity* (from *university*) and *phone* (from *telephone*). Second, there is 'middle clipping', that is, deleting both the first and last part and keeping the middle part, which is rare, for example, *jams* (from *pyjamas*) and *flu* (from *influenza*). Finally, there is 'back clipping', that is, deleting the second part and keeping the first part, for example, *exam* (from *examination*), *brill* (from *brilliant*), *ad* (from *advertisement*). It is widely used with names, for example, Max (Maximilian) and Rich (Richard).

Clipping may interact with compounding. Compounds can be clipped, for example, *pub* from *public house*; and compounds can be created from clipped words, for

example, *hi-fi* (from *high fidelity*), *sci-fi* (from *science fiction*), *Britcom* (from *British comedy* on TV and film, e.g., *Blackadder*).

Hypocorisms

Hypocorism is used to refer to words formed by suffixing a vowel, usually *–y* or *–ie* [i] to a monosyllabic root or by suffixing *–y* or *–ie* [i] after clipping has reduced a longer simplex or compound word to one syllable. It is used to create the familiar forms of names, for example, *Johnnie* (from John), *Vicky* (Victoria) and *Mandy* (Amanda). It is also used for common nouns, for example, *chippy*, *movie*, *kiddy*, *biccy*, *brownie* and *bookie*. These contractions are usually referred to as **diminutives**. But this label is not always appropriate, especially in Australian English where this type of word-formation is most widely used. As well as being used in a diminutive sense, hypocorisms are used in Australian English for wordplay and for indicating empathy (Simpson, 2001). Examples include *baggie* (large school bag), *barrie* (as in 'give someone a barrie', i.e., give someone a ride on the bar of your bike), *Brizzie* (Brisbane), *Chrissy* (Christmas) and *Saffie* (a South African person). Another vowel commonly used in hypocoristics is *–o* (cf. *doco* (documentary), *journo* (journalist), *Nasho* (National Service)). Particularly interesting is the case of *hottie* (hot water bottle in Australian, and British slang). But in British slang *hottie* has been apparently re-analysed as a blend of *hot* + *(cu)tie* → *hottie*. Original *hotties* were all female. Subsequent widening of meaning made *hottie* a unisex adjective, applicable to women and men, prompting Safire to observe in the *New York Times Magazine* , '[R]arely do we come across a word that can be applied, with lust aforethought, to either sex.' This usage has spread to the other side of the Atlantic.

Acronyms and abbreviations

Word contraction is taken to its logical conclusion in **acronyms** and **abbreviations** (also called 'initialisms'). In this type of word-formation a group of words representing a concept or the name of an organization is reduced to their initial letters, which are then treated as a word. In the case of an abbreviation, the reduced form does not result in well-formed syllables and so cannot be pronounced as a word. Rather, the letters are sounded out independently, for example, *EU* (European Union), *BBC* (British Broadcasting Corporation), *RBS* (Royal Bank of Scotland).

In the case of acronyms, contraction delivers initial letters that constitute well-formed syllables and the string forms a perfectly normal word, for example, *NATO* (North Atlantic Treaty Organization), *NICE* (National Institute for Clinical Excellence), *laser* (light amplification by the stimulated emission of radiation), radar (radio detection and ranging), *sim (card)* (Subscriber Identity Module (card)).

Borrowing

As well as using its own resources to enrich its lexicon, English very readily incorporates words from other languages into its vocabulary. Words like *resources*, *lexicon*, *incorporates*, *languages* and *vocabulary* are foreign imports. The role of borrowing is explored in Chapter 14.

4.8 Summing up

This chapter has presented an overview of word structure and patterns of word-formation. Defining a word is a knotty problem. It is best approached by distinguishing various senses in which the term 'word' is used (lexeme,

grammatical word and word-form). With regard to structure, some words are simple and others complex. Normally, complex words can be decomposed into smaller structural and semantic units which are called morphemes. Many morphemes have variants which are called allomorphs that represent them in various contexts. Morphemes are classified as roots or affixes; free or bound.

Morphology has two broad types of affixes, namely inflectional affixes and derivational affixes. The former are there to ensure that the word has the right form if it occurs in a given syntactic slot while the latter serve to create new lexemes from bases. Various morphological operations are used in English derivational morphology, namely affixation, conversion and compounding. In inflectional morphology, too, there are a number of processes that can be identified, including affixation, internal vowel change and suppletion. There are also various additional ways of creating lexemes: coinage, eponymy, backformation, blending, clipping, hypocorism, acronyms and abbreviations (and borrowing).

A thread that runs through the account of English word-formation presented here is the importance of seeing morphology in a wider context. Morphology interacts with meaning, syntax and phonology in intricate ways. So, word-formation can't be undertaken without considering the mutual relationship of morphology with phonology, with syntax, and with semantics. The next chapter investigates words from a syntactic perspective and the semantics chapter (chapter 9) looks at words from a meaning perspective. The study of words shows clearly the fact that there is no semi-detached component of grammar where life is lived by linguistic entities in total isolation from the rest of grammar.

Recommended readings

Accessible introductions to word-formation in English include Carstairs-McCarthy (2018), Bauer (2003), Katamba (2005) and Stockwell and Minkova (2001), which includes a useful historical dimension. For more advanced general textbooks, see Matthews (1991), and Aronoff and Fudeman (2005). Marchand (1969) is the classic reference work on English word-formation.

Chapter 5

Grammar: Words (and Phrases)

Geoffrey Leech

5.1 Introducing word classes

Everyone knows that any piece of language, such as a written text or a spoken piece of dialogue, consists of **words**. But how many kinds of words are there? This is a matter for discussion and argument, but for the purposes of this chapter there will be 11 **word classes** (i.e., word classes which are not part of other word classes) – commonly known as 'parts of speech'. By the end of this and the following chapter, you should be able to take a piece of language, such as a paragraph of text, and label each word as one of these 11 word classes and, even in uncertain cases, you will probably be able to make a good guess. Words are the main theme of this chapter – but we cannot talk seriously about words without talking about the 'word-chunks' that form larger units – **phrases** – a topic we look into more fully in the next chapter. Also, towards the end of this chapter, we will examine critically, in arriving at those 11 types, an assumption we have made about word classes.