Chapter 4: Morphology   Hans-Jörg Schmid

Key terms: morpheme, allomorph, inflection, analytic languages, synthetic languages, derivation, word-formation, compounding, prefixation, suffixation, conversion, back-formation, blending, clipping, acronym

Preview: In this chapter you will first learn to segment words into their smallest meaningful parts, their morphemes. Different types of morphemes will then be distinguished on a number of dimensions. The classes arising from such distinctions are useful because they allow us to formulate generalizations about the properties shared by the members of these classes and the restrictions they are subject to. A second type of generalization covered in this chapter concerns the patterns and rules which underlie the formation of complex lexemes, i.e. words that are made up of more than two lexical morphemes. This is the realm of word-formation. You will be introduced to the range of word-formation patterns that can be used to form new words with the help of existing words and morphemes, including compounding, i.e. the joining of two or more words to form a new complex lexeme (e.g. *interest rate, washing-machine* or *watertight*), prefixation, yielding words such as ***dis**agree,* ***un**just* or ***ex**-minister*, and suffixation (e.g. *agree**ment**, just**ify**, minis**terial**)*. Further word-formation patterns, which are less regular and transparent, include conversion (*hammer* ₙ → *to hammer* ᵥ or *empty* ₐdⱼ → *to empty* ᵥ), back-formation (e.g. *to sightsee ← sightseeing; to burgle ← burglar*), blending (e.g. *infotainment ← information + entertainment*), clipping (e.g. *ad ←advertisement, phone ← telephone*), and the formation of acronyms or initialisms from fixed sequences of words (*URL ← unique resource locator*; *NATO ← North Atlantic Treaty Organization*).

4.1 Introduction

Generally speaking, the linguistic discipline of morphology – the term is derived from the Greek word *morphos* meaning 'form' – examines the internal makeup and structure of words as well as the patterns and principles underlying their composition. In doing so, morphology straddles the traditional boundary between grammar (i.e. the rule-based, productive component of a language) and the lexicon (i.e. the idiosyncratic, rote-learned component). Morphology looks at both sides of linguistic signs, i.e. at the form and the meaning, combining the two perspectives in order to analyse and describe both the component parts of words and the principles underlying the composition of words.

Unlike phonology, morphology does not analyse words in terms of syllables but in terms of **morphemes**, i.e. components of words that are carriers of meanings. For example, while the words *father* and *teacher* both consist of two syllables, *father* represents only one morpheme (meaning 'male parent'), whereas *teacher* consists of two: the verb *teach* ('instruct') and the nominalizing suffix *-er* ('someone who does something'). The most frequently found definition of the notion of morpheme states that it is the 'smallest meaning-bearing unit' in a given language. As the example of *father* has shown, morphemes can coincide with simple words, or more precisely, **simple lexemes**, i.e. abstract representations of words uniting forms and (bundles of related) meanings, but they can also constitute parts of **complex lexemes**, which are in turn defined as lexemes consisting of more than one morpheme. Unlike *father*, then, *teacher* is an example of a complex lexeme.

The study of morphology is traditionally divided into two major areas. The first is known as **inflectional morphology** and deals with the markers of grammatical categories such as CASE, NUMBER, TENSE and ASPECT. These inflectional morphemes are attached to lexical stems and create word-forms (rather than new words). For example, the verb *employ* can occur in the base-form *employ* when no inflectional morpheme is added, in the form *employs* when the morpheme marking agreement with a third person singular subject is attached, in the form *employed* when marked by the past tense or the part participle morpheme, and in the *ing*-form *employing*, used, among other things, for encoding the progressive aspect. The second major branch of morphology is **word-formation**, whose scope includes the direct terminological counterpart to inflectional morphology, **derivational morphology**, but goes beyond that. The field of word-formation deals with the patterns and rules guiding the formation of new words (rather than just word-forms of existing words). From this perspective, the word *unemployment*, for instance, would first be segmented into the base *employ* and the derivational morphemes *un-* and *-ment*, and it would be stated that the affixes *un-* and *-ment* are added to the base *employ*, thus manifesting the word-formation types of **prefixation** and **suffixation** respectively. In addition to derivational morphology, word-formation encompasses the study of **compounding** (e.g. *employment agency*) and also those word-formation types that do not use morphemes as their basic building-blocks, i.e. non-morphemic types such as **blending** (e.g. *infotainment* ⬅ *information* and *entertainment*) and **clipping** (e.g. *flu* ⬅ *influenza*). The scope of the fields of morphology and word-formation is summarized in Figure 4.1.

In appreciating this figure you should keep in mind, however, that the situation is not quite as simple and clear as suggested by the neat division: firstly, the word-formation type of compounding does not really fall within the scope of derivational morphology, but is placed in the same branch as prefixation and suffixation because it shares with these the property that it uses morphemes as basic building blocks. Secondly, **conversion**, i.e. the transfer of a word from one word class to another without the addition of a morpheme, and **back-formation**, as in the verb *to sightsee* derived from the longer noun *sightseeing*, could be seen as relying on morphemes, too, but this is much less straightforward than is the case in the other types of morphemic word-formation patterns.

Figure 4.1: The scope of morphology and word-formation (adapted from Schmid 2011: 15)
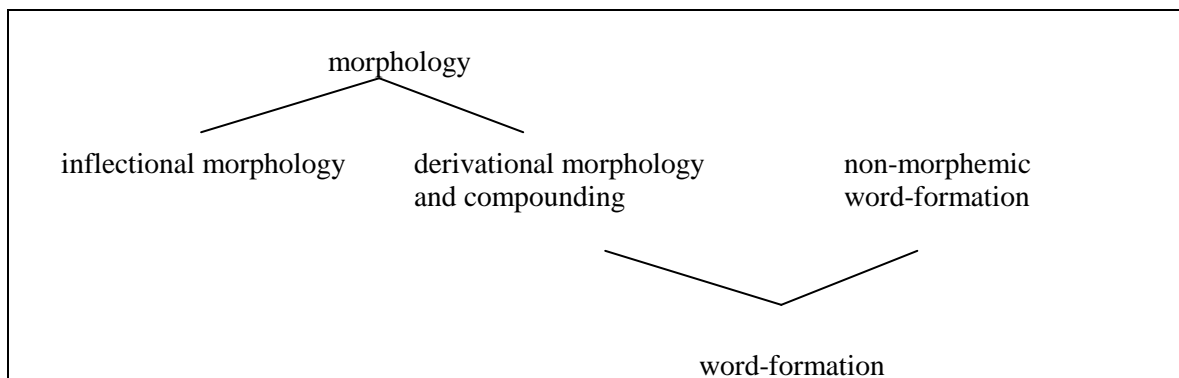
4.2 Morphemes and other morphological building-blocks

As has been pointed out above, morphemes are defined as smallest meaning-bearing units. Morphemes can be classified in various ways. One common classification you have already learnt about above separates those morphemes that mark the grammatical forms of words (*-s, -ed, -ing* and others) from those that form new lexemes conveying new meanings, e.g. *un-* and *-ment*. The former morphemes are **inflectional morphemes** and form a key part of grammar, the latter are **derivational morphemes** and play a role in word-formation, as we have seen. The following criteria help you to distinguish the two types:

- Effect: Inflectional morphemes encode grammatical categories and relations, thus marking word-forms, while derivational morphemes create new lexemes.

- Position: Derivational morphemes are closer to the stem than inflectional morphemes, cf. *amendments* ($amend_{\text{stem}} - ment_{\text{derivational}} - s_{\text{inflectional}}$) and *legalized* ($legal_{\text{stem}} - ize_{\text{derivational}} - ed_{\text{ inflectional}}$).

- Productivity: Inflectional morphemes are highly productive, which means that they can be attached to the vast majority of the members of a given class (say, verbs, nouns or adjectives), whereas derivational morphemes tend to be more restricted with regard to their scope of application. For example, the past morpheme can in principle be attached to all verbs; suffixation by means of the adjective-forming derivational morpheme *-able*, however, is largely restricted to dynamic transitive verbs, which excludes formations such as **bleedable* or **lieable*.

- Class properties: Inflectional morphemes make up a closed and fairly stable class of items which can be listed exhaustively, while derivational morphemes tend to be much more numerable and more open to changes in their inventory.

Both inflectional and derivational morphemes must be attached to other morphemes; they cannot occur by themselves, in isolation, and are therefore known as **bound morphemes**. **Free morphemes**, on the other hand, are autonomous, can occur on their own and are thus also words at the same time. Technically, bound morphemes and free morphemes are said to differ in terms of their 'distribution' or 'freedom of occurrence'. As a rule, lexemes consist of at least one free morpheme.

A third way of classifying morphemes relies on the kinds of meanings they encode. **Grammatical morphemes** serve the purpose of signalling grammatical categories and encoding relational meanings, while **lexical morphemes** carry richer conceptual, more autonomous meanings. Note that this distinction overlaps partly, but not fully, with the one between inflectional and derivational morphemes. In fact, as shown in Table 1, inflectional morphemes form the subclass of bound grammatical morphemes, whereas derivational morphemes are bound lexical morphemes.

Table 4.1 gives a survey of a widespread way of classifying morphemes in terms of a cross-tabulation of the dimension of distribution/freedom of occurrence (free vs. bound) and meaning (lexical vs. grammatical).

Table 4.1: A cross-classification of types of morphemes

	lexical morphemes	grammatical morphemes
free morphemes	= content words (e.g. *paper, slim, run*) • semantically and distributionally more autonomous • can be inflected • rich conceptual content	= function words (e.g. *to, the, of*) • semantically and distributionally less autonomous • cannot be inflected • mark grammatical relations
bound morphemes	= derivational morphemes (e.g. *re-, -ize, -able*) • create new lexemes • closer to the stem • more restricted productivity • more open class	= inflectional morphemes (e.g. *-s, -ed, -est*) • mark word-forms • more distant from the stem • highly productive • closed class

The table also indicates that the class of free grammatical morphemes contains so-called function words such as *the, of* or *to*, which mark grammatical relations, cannot be inflected and are semantically and distributionally much more restricted than free lexical morphemes (i.e. so-called content words). Content words belong to the word-classes of nouns, adjectives and adverbs and form the large majority of verbs, while function words comprise articles, conjunctions, prepositions and particles as well as the so-called primary verbs *be, have* and *do*, which contribute to the encoding of grammatical categories such as TENSE and ASPECT (*I **have been** running*), NEGATION (*She **does** not eat shrimp.*), VOICE (*He **was** scratched by the dog*) or sentence MOOD (*Does she eat garlic?*).

While the distinctions introduced so far seem straightforward enough, it turns out that implementing the definition of morphemes as *smallest meaning-bearing components of words* is not an easy task. One complication arises from the fact that short and seemingly simple word forms can express sets of meanings which are encoded by several morphemes in other words. Consider, for example, the form *sang* carrying the meanings of the lexical morpheme {sing} and the grammatical morpheme {past}, which are expressed by two morphemes in *shouted*, *kissed* and many other verbs.

Secondly, as will be discussed in greater detail in Section 4.3, morphemes are not always realized by the same form but by a number of variants, so-called **allomorphs**, depending on the environment in which they occur. This is particularly relevant for inflectional morphemes. The form *sang* mentioned above can in fact be treated as a rather unpredictable allomorph of the {past} morpheme. More regular allomorphs can be

identified in the forms *smiled*, *laughed* and *greeted*, where the past morpheme is realized by the allomorphs /d/, /t/ and /ɪd/ respectively.

Thirdly, you can face difficulties when trying to segment words into morphemes because a seemingly reasonable formal analysis is not matched by a semantic one, or because your segmentation does not leave you with a free morpheme, as is usually required. The word *refer* is a case in point. You may well be inclined to divide this word into the morphemes *fer*, which you also find in *transfer, infer, confer* and *prefer*, and the derivational prefix *re-* occurring in large numbers of other verbs. What you soon realize, however, is that neither of these two potential morphemes is free, and that you will not find it easy to work out a meaning for the form *fer* which is shared by all the verbs in which it occurs (unless you happen to know that it is derived from Latin *ferre* 'to carry', but even then things do not quite make sense). Many of these cases have to do with the fact that English borrowed large numbers of words from Latin which were already prefixed and suffixed in that language, but did not bother to borrow the bases – cf., e.g., *describe, inscribe, subscribe, prescribe* but **scribe* (as a verb) or *insist, desist, consist, persist, resist* but **sist*. To solve this analytical dilemma, in some accounts of morphology (e.g. Stockwell and Minkowa 2001: 61–62) the bases of these lists of forms are given the special status of bound roots, which can be considered as somewhat untypical kinds of lexical morphemes.

Fourthly and finally, analytical problems arise because some forms can be put to use as both lexical and grammatical morphemes. The form *-ing*, for example, functions as a grammatical, inflectional morpheme participating in the formation of progressives (*she was knocking on his door*) and as a lexical, derivational morpheme forming adjectives from verbs (*interesting, exciting*) or nouns from verbs (*meeting, building*). In this case you could argue that the two functions are closely related and that the morpheme has several similar meanings. You could say that the morpheme is polysemous. In contrast, the use of *-er* as a nominalizing derivational suffix (as in *teacher*) is clearly unrelated to its use in the formation of the comparatives of adjectives (*wider, rougher*, etc.). Two different morphemes happen to have the same form, which is a case of homonymy rather than polysemy.

KEY POINTS: Morphemes

- morphemes are the smallest meaning-carrying units of a language
- simple lexemes consist of one morpheme only, while complex lexemes have at least two lexical morphemes
- inflectional morphemes can be distinguished from derivational morphemes on the basis of their effects on the base (marking of word-forms vs. creation of new lexemes), their position vis-à-vis the stem (more distant vs. closer), their productivity (highly productive vs. restricted) and their class-properties (closed class vs. open class)
- free morphemes are autonomous, while bound morphemes cannot occur in isolation
- grammatical morphemes mark grammatical categories and relations, lexical morphemes carry conceptual meanings

---

Exercise 4.1

Using the information provided in this section and keeping in mind the four complications, you can now tackle the task of segmenting the following passage into morphemes and classifying them along the lines summarized in Table 4.1. Follow the format suggested below the text. (List of abbreviations: gr = grammatical; lex = lexical; fr = free; bd = bound).

While his granddaughters were still playing with their laptops and desktop computers in the living-room, Granddad found the necessary picnic supplies in the fridge and began to take them to his beloved flashy BMW convertible.

gr	gr	lex	lex		gr	...
while	*his*	*grand*	*daughter*		*s*	...
fr	fr	fr	fr		bd	…

4.3 Inflectional morphology

Languages differ considerably with regard to the extent to which they employ inflectional morphemes to mark grammatical categories and the way in which these morphemes are combined. On one end of a continuum are **analytic languages**, which do not signal

grammatical categories and relations by means of inflectional morphemes but instead by other strategies such as fixed word order, auxiliaries and particles. As present-day English can muster only a relatively small number of inflectional morphemes, it comes quite close to acting as a representative of such a language. The full inventory of bound grammatical morphemes, which is listed in Table 4.2, amounts to less than ten items:

Table 4.2: Inflectional morphemes in present-day English

word-class	morpheme	functions/meanings	grammatical category
noun	{plural}	marking of plural	NUMBER
	{genitive}	marking of genitive, possession, part-of, etc.	CASE
verb	{3rd person}	3rd person singular present, marking agreement with subject	PERSON, AGREEMENT
	{ing}	present participle, marking of progressive aspect	ASPECT
	{ed$_1$}	simple past	TENSE
	{ed$_2$}	past participle, used for present perfect and passive voice	TENSE
adjective	{er}	comparative	GRADATION
	{est}	superlative	

Controversial further candidates are the form {-th}, which forms ordinal numbers (*fourth, fifth*) and the adverb-forming suffix {-ly}, which is sometimes treated as a derivational morpheme because it causes a change of word-class, and sometimes as an inflectional suffix because its productivity is almost unrestricted and the changes in meaning and grammatical function are very limited.

Synthetic languages, on the other hand, encode large numbers of grammatical categories by attaching inflectional morphemes. German, for example, differs substantially from English in this respect, as it can mark four CASES (nominative, accusative, genitive and dative), NUMBER (singular and plural) as well as the GE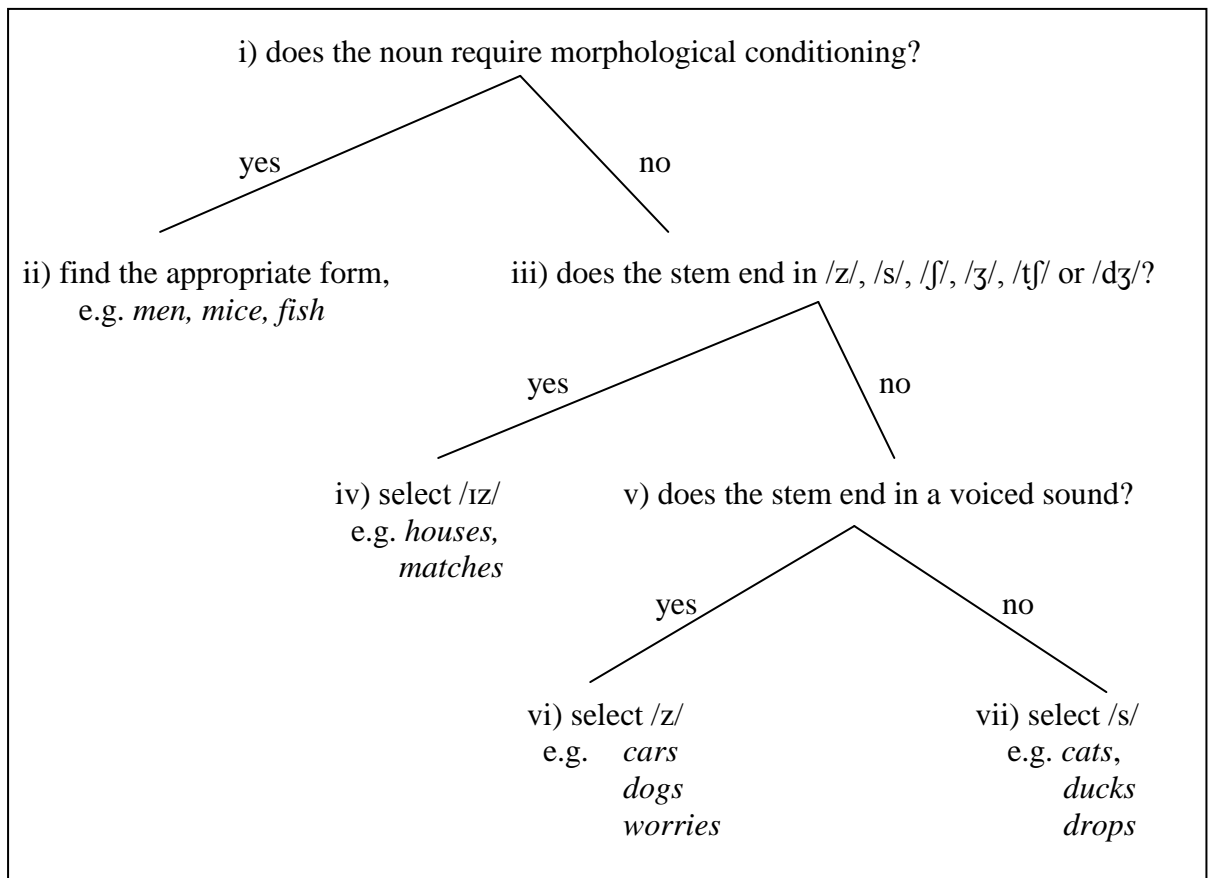NDER of nouns (masculine, feminine and neuter). Adjectives can be marked for CASE and NUMBER, and the type of marking differs depending on whether the weak or strong declension is required (cf. *ein gutes Buch* 'a good book' vs. *das gute Buch* 'the good book'). Verbs are marked not only for TENSE by inflectional morphemes, but also for PERSON, NUMBER and MOOD (indicative vs. conjunctive). In addition, articles and pronouns are important markers of CASE, PERSON, GENDER and NUMBER. While many of these distinctions have collapsed into the same forms (a phenomenon known as syncretism), there

can be no doubt that the grammar of German relies on inflectional marking to a much greater extent than English does. Languages which are even richer in inflectional markers, especially markers of much larger numbers of CASES, are by no means uncommon, both within the branch of Indo-European languages and elsewhere. Russian, for example, has inflectional markers for as many as six cases, Hungarian for more than a dozen.

Within the group of synthetic languages, we can distinguish between so-called **fusional languages** like German or Latin, which often express a whole set of grammatical meanings in one form – as is the case in the Latin *bonus* ('good') where the form *-us* encodes the morphemes 'masculine', 'singular' and 'nominative' – and **agglutinating languages**. In these languages, examples of which include Turkish, Finnish, Mongolian and Japanese, word-forms and even phrases and clauses are produced by joining sequences of morphemes that neither overlap nor collapse into forms expressing several meanings (cf. the Turkish form *evlerinizin* 'of your houses', joining the morphemes *ev* 'house', *ler* 'plural', *in* '2nd person poss. pron.', *iz* 'plural' and *in* 'of').

In the course of its historical transition from Old English to Modern English, the English language has undergone a shift from more synthetic to more analytic. This entails that it has not only lost most of its inflectional morphemes, but has also become more 'regular' by leaving behind a massive number of allomorphs. Almost the entire system of differently conjugated classes of strong verbs and differently declined classes of nouns which were once typical of Germanic languages has disappeared. Irregular forms such as the past tense forms *sang, stood* or *told* and the plural forms *mice* and *oxen* are indeed only small remnants of the complex system of former allomorphs. Today, the choice of the remaining variants can be determined either by the stem as such (**morphological conditioning**) or by the final sound of the stem to which an inflectional morpheme is added (**phonological conditioning**). All the 'irregular' forms cited above are morphologically conditioned, since the stems *sing*, *stand*, *tell*, *mouse* and *ox* are responsible for how they are modified in order to mark the past or plural. Phonological conditioning is less idiosyncratic and can be captured in terms of a small set of rules, which are illustrated for the plural morpheme in Figure 4.2:

Figure 4.2: Rules of allomorphy illustrated for the English plural morpheme

i) does the noun require morphological conditioning?

yes

no

ii) find the appropriate form,
e.g. *men, mice, fish*

iii) does the stem end in /z/, /s/, /ʃ/, /ʒ/, /tʃ/ or /dʒ/?

yes

no

iv) select /ɪz/
e.g. *houses,
matches*

v) does the stem end in a voiced sound?

yes

no

vi) select /z/
e.g.　*cars
dogs
worries*

vii) select /s/
e.g. *cats,
ducks
drops*

The allomorph systems of other languages are again much more complex. Again taking German as an example, Table 4.3 lists the more systematic allomorphs of the plural morpheme:

Table 4.3: Extract of the system of German plural allomorphy

Declination and allomorph	gender	examples without umlaut	examples with umlaut
N-declension: *-n* or *–en*	masc. fem. neutr.	*der Bär/die Bären* 'the bear/bears' *die Pfeife/die Pfeifen* 'the pipe/pipes' *das Auge/die Augen* 'the eye/eyes'	-
E-declension: *-e*	masc. fem. neutr.	*der Stein/die Steine* 'the rock/rocks' *die Erkenntnis/die Erkenntnisse* 'the insight/insights' *das Rohr/die Rohre* 'the tube/tubes'	*der Hut/dieHüte* 'the hut/huts' *die Maus/die Mäuse* 'the mouse/mice' -
R-declension: *-er*	masc. neutr.	*derLeib/die Leiber* 'the body/bodies' *das Bild/Bilder* 'the picture/pictures'	*der Mann/die Männer* 'the man/men' *das Glas/die Gläser* 'the glass/glasses'
zero-declension: -Ø	masc. neutr.	*der Rahmen/die Rahmen* 'the frame/frames' *das Fenster/die Fenster* 'the window/windows'	*der Boden/die Böden* 'the floor/floors' -
S-declension *-s*	masc. fem. neutr.	*derPark/die Parks* 'the park/parks' *die Bar/die Bars* 'the bar/bars' *das Auto/die Autos* 'the car/cars'	-

A wide variety of further forms and rules have to be used for other nouns depending on an intricate combination of factors including meaning, derivational suffixes, origin (native or foreign) and others. Arguably, it is this comparative morphological complexity and unruliness

of German that has caused judgments of the type "Life is too short to learn German", found on T-shirts and attributed variously to Richard Porson, Mark Twain and Oscar Wilde.

Key Points

- synthetic languages have a large number of inflectional morphemes, while analytic languages can only muster few of them
- German is located further towards the synthetic end of the continuum than present-day English, which has gradually become more analytic over the past fifteen centuries
- among the synthetic languages, fusional languages like Latin often encode several grammatical meanings in one form, while agglutinating languages such as Turkish join sequences of morphemes
- inflectional morphemes have allomorphs whose forms can depend on the stem (i.e. be morphologically conditioned) or on the final sound of the stem (i.e. be phonologically conditioned)

Exercise 4.2

Figure 4.2 has provided a maximally systematic and economical way of describing the allomorphy of the English plural morpheme. Try to transfer the logic behind this figure to the English past tense morpheme and produce a similar figure. You can use the following dataset as a basis for your classification: *kissed, loved, watched, smiled, sat, put, sang, laughed, was, cost, ruined, rated, rode, pleaded*. Make sure you pay attention to how the sounds represented by *-ed* are pronounced. For example, in *kissed* the ending *-ed* is realized by the sound /t/, in *loved* by the sound /d/, and in *rated* by the sounds /ɪd/. Try to determine how these choices are conditioned. Irregular forms (e.g. *sang, was*) can also be treated in analogy to Figure 4.2.

4.4 An analytical and descriptive system for the study of word-formation

Before we can look at the different types of word-formation patterns in more detail, it will be helpful to introduce a diagnostic system (cf. Schmid 2011: 95ff). This essentially serves two main functions, one related to the analysis of existing words and one related to the way in which new words are created. Obviously, these two functions are intertwined. Firstly, the system helps to reveal similarities and differences in the body of existing complex words; here it has an analytical and descriptive function. Secondly, we can use the information gained from such analyses to formulate general ideas concerning how new words can be

formed. The rationale behind this is that the way in which existing words have been formed will not be too different from the way in which new words can and will be formed. Irrespective of the format in which generalizations are stated, e.g. as patterns, rules or schemas, they essentially have the function of capturing speakers' tacit knowledge about how to decompose and form words. This means that the system to some extent reflects some sort of 'grammar' of word-formation, Note that since the system ultimately aims at the description of regularities, it is more applicable to the more regular and predictable field of morphemic word-formation than to that of the more haphazard and unpredictable non-morphemic patterns.

Morphological form: The first level in the proposed system concerns the analysis and description of word-formation products in terms of their morphological form or shape. This is done by segmenting complex lexemes into their morphemes and other potentially meaning-bearing constituents, and classifying them. The description of the morphological form can either be presented in terms of the morpheme classification introduced in Section 4.2, or can make use of the terms *base, affix, prefix, suffix*, etc. It is helpful for further analytical steps to mark the word-classes of free lexical morphemes and the word-class changes caused by bound lexical morphemes, especially suffixes. Table 4.4 provides illustrations of the terminology and the kinds of elements that can be used for this part of the analysis. The examples will be taken up for further discussion below.

Table 4.4: Illustrating the analysis in terms of morphological form

disappointment	*dis-* lexical bound prefix	*appoint* lexical free base$_V$	*-ment* lexical bound suffix$_N$	
unemployment	*un-* lexical bound prefix	*employ* lexical free base$_V$	*-ment* lexical bound suffix$_N$	
armchair	*arm* N	*chair* N		
paperback	*paper* N	*back* N		
trade union leader	*trade* lexical free N	*union* lexical free N	*lead* lexical free base$_V$	*-er* lexical bound suffix$_N$

Morphological structure: The descriptions in Table 4.4 fall short of accomplishing the mission of teasing apart things that look similar but are in fact different. Consider the two examples *disappointment* and *unemployment*. Table 4.4 renders identical analyses in terms of morphological forms for these two nouns. However, these analyses conceal the fact that the

two nouns differ with regard to their internal constituent structure, i.e. with regard to the question as to which elements belong together more closely than others. Looking at *unemployment* first, you will realize that here the suffix belongs more closely to the base than the prefix does, because the verb **to unemploy* which would have to serve as a base for the suffixation does not exist. The formation history, so to speak, must therefore be *employ* → *employment* → *unemployment*. This means, as is shown in Figure 4.3, that *employ* and *-ment* are so-called immediate constituents. For *disappointment*, the situation is different. Here it is much more likely that the prefix was added to the base first, yielding the verb *to disappoint*, with the suffix being added in a second step, since the noun *appointment* does not seem to be semantically related to *disappointment*. The morphological structures of the two nouns thus differ, as is shown in Figure 4.3.

Figure 4.3: Comparing the morphological structures of *unemployment* and *disappointment*
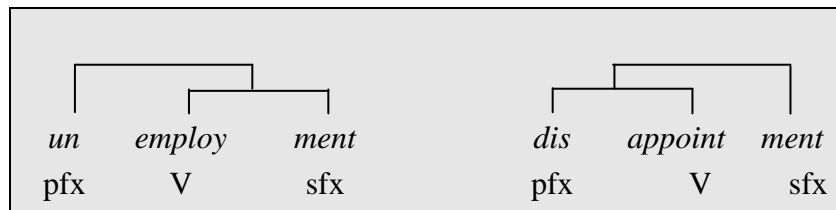

Adequate descriptions of the morphological structures of complex lexemes require even more information, however. The two compounds *armchair* and *paperback*, which are identical in terms of morphological form and immediate constituents, lend themselves to an illustration of this aspect. In the case of *armchair*, the first constituent *arm* modifies the second constituent *chair*. Both grammatically and semantically, *chair* can be considered the **head** of the compound, while *arm* functions as a **modifier**. A suitable paraphrase of the meaning of this compound could begin with the head and add the extra information provided in the modifier: 'a chair that has arms'. An analogous paraphrase is clearly impossible for *paperback*, since the meaning of this word is certainly not 'a back that is made from paper' but rather 'a book that has a back made from paper'. This indicates that the head of the compound *paperback* is not *back* but could be *book* despite the fact that this is not part of the morphological form of *paperback*. The exemplary comparison of the two modifier-head structure reveals that *armchair* and *paperback* are not two of a kind and should therefore not be lumped together.

The distinction between modifier and head is an important general descriptive principle in word-formation, which, just like the idea of immediate constituents, has been taken over from

syntax. As in syntactic structures, in English it is generally the case that in complex lexemes modifiers also precede heads. Heads are therefore the right-most constituents of complex lexemes and determine their word-classes. In the fairly complex formation *trade union leader* it is the last constituent *-er* which marks the whole unit as a noun. Furthermore, again as in the syntactic analysis of sentences, in the morphological analysis of complex lexemes we generally strive for a binary, i.e. two-way branching of constructions into immediate constituents. Lexemes consisting of more than three morphemes can usually be accounted for by several hierarchical layers of binary modifier-head combinations. This is shown in the exemplary analysis given in Figure 4.4, which demonstrates the whole scope of the analysis in terms of morphological form and structure:

Figure 4.4: Illustration of description in terms of morphological form and structure: *trade union leader*
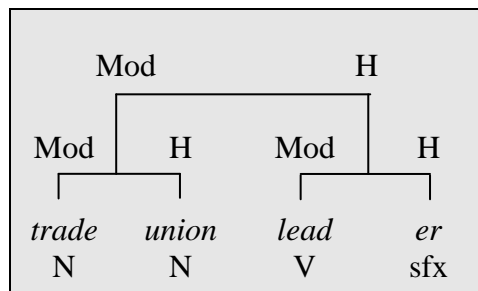
While the analytical steps described so far already go a long way towards an adequate account of complex lexemes, they do not yet tell the whole story. What is needed to obtain the full picture is a description of the internal semantic structure of complex lexemes, including the semantic relations between the constituents. Consider as a first illustration the compounds given in (1), all of which have the noun *chair* as head. Using the system set up so far, you would be able to come up with a number of interesting observations: that all of these compounds consist of two free lexical morphemes; that they all represent a modifier-head structure; and that *high chair* and *swivel chair* differ from the rest of the group in that they have an adjective and a verb respectively as modifiers, rather than a noun. This account would miss out on important further differences, however, which concern the semantic relations linking the constituents of these compounds. As we have seen, the meaning of *armchair* can be paraphrased as 'a chair that has arms', indicating that the relation between *arm* and *chair* can be described as a possessive or part-whole one. As is pointed out in (1), however, the other compounds in the list encode distinctly different semantic relations, including comparison, identity and others:

(1)

armchair:	'a chair that has arms'	→	POSSESSION/PART-WHOLE
barrel chair:	'a chair that is shaped like a barrel'	→	COMPARISON
bedchair:	'a chair that be turned into/is also a bed'	→	IDENTITY
cane chair:	'a chair made from cane'	→	SUBSTANCE, MATERIAL
deck chair:	'a chair that is found on the deck of ship'	→	LOCATION
high chair:	'a chair that is unusually high'	→	SIZE
swivel chair:	'a chair that allows you to swivel'	→	FUNCTION

A similar problem arises in the analysis of suffixations and prefixations, where form-meaning ambiguities are also very common. The noun *declaration* is a good example of the widespread phenomenon that nominalizations can highlight different aspects of activities, processes and states:

(2)

declaration   'action of declaring something', e.g. *his declaration took two hours*

'result of declaring something', e.g. *I did not believe his declaration*

'product of declaring something', e.g. *they signed a declaration*

The relations between prefixes and their bases can also vary considerably, even when the same prefix is used. What this shows is that an analysis and description of the internal semantic structures and relations must complement that of the morphological forms and structures.

KEY POINTS: Analysis and description of word-formation types

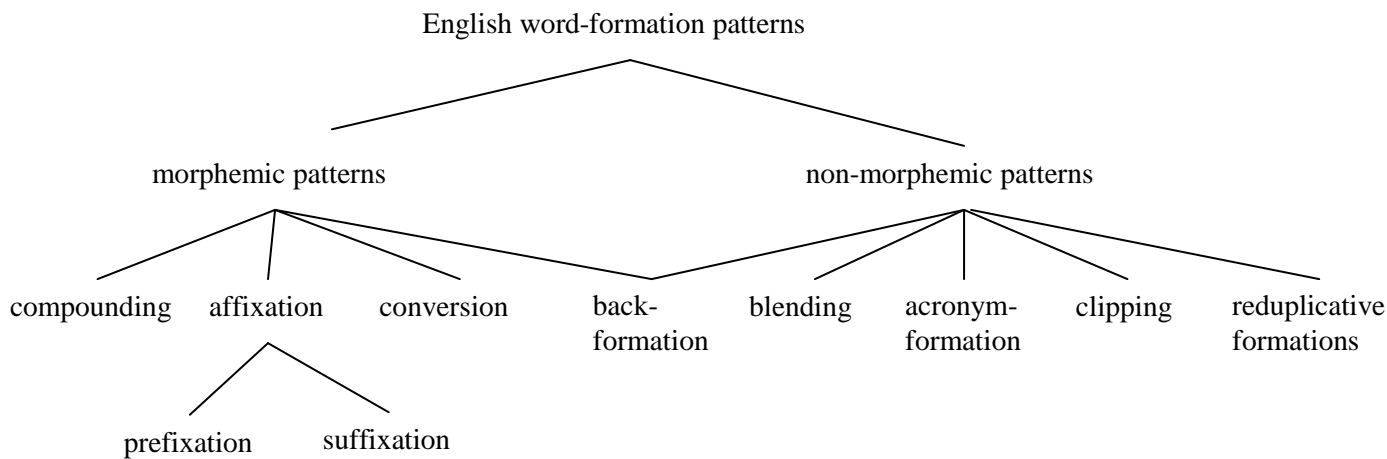An adequate account of the structure of complex lexemes should consist of information on

- their morphological forms,
- their morphological structures in terms of immediate constituents and modifier-head relations,
- and their internal semantic structures and relations.

This now puts us in a position to have a systematic look at the system of word-formation in English, beginning with a general survey of the basic patterns.

4.5 Survey of English word-formation patterns

Figure 4.5 renders the most common way of classifying word-formation patterns in English and is to serve as a frame for the following sections (cf. Plag 2003, Lieber 2005, Schmid 2011).

Figure 4.5: Survey of basic word-formation patterns in English

English word-formation patterns

morphemic patterns   non-morphemic patterns

compounding   affixation   conversion   back-formation   blending   acronym-formation   clipping   reduplicative formations

prefixation   suffixation

4.5.1 Compounding

Compounding is the process of joining at least two free lexical morphemes or simple lexemes to form a complex lexeme. Compounds can be distinguished from syntactic phrases with the help of a number of criteria: in compounds, the main stress is typically, but not always, on the first constituents (cf. *'blackbird* vs. *black 'bird*); the first constituent cannot be inflected (cf. **wallspaper*); the head cannot be replaced by *one* in coordination (cf. **let's buy a newspaper and a wall one*); and compounds are typically lexicalized, which means that the meaning of the compound tends to differ and go beyond the meanings of its parts. For example, a *holiday* is not just, or no longer, a *holy day*, as its components suggest, but typically extends over several days and mainly marked by the fact that people do not work.

In terms of morphological form, if compounds consist of nothing else than two lexemes (e.g. *wallpaper, mousemat, daydream*), they are commonly called **root compounds** and described in terms of the word-classes of their constituents (cf. Table 4.5). Compounds containing bound morphemes in addition to free ones, e.g. *meeting point, theatre-goer* or *good-looking*, are known as **synthetic compounds**.

The most frequent and productive types of compounds are listed in Table 4.5 (Schmid 2011: 122):

Table 4.5: Most frequent and productive types of compounds in terms of morphological form

---

Nominal compounds:

N + N: *backbone*, *barman*, *nutshell*, *pony tail*, *seat-belt*, *timetable*, *wallpaper*
Adj + N: *greenhouse*, *high-chair*, *smalltalk*, *stronghold*
[V + ing] + N: *dancing girl*, *building-block*, *dressing gown*, *racing car*
V + N: *cease-fire*, *copyright*, *showroom*, *stopgap*
N + [V + ing]: *credit rating*

Adjectival compounds:

N + Adj: *accident-prone*, *carefree*
N + [V + ing]: *awe-inspiring*, *eye-catching*, *time-consuming*
Adj + [V + ing]: *good-looking*, *hard-drinking*

True verbal compounds, i.e. complex lexemes with verbs as heads which are actually the result of a compounding process, do not seem to exist in English or are in any case very rare. Potential candidates such as *to sightsee*, *to babysit* or *to handmake* are not verbal compounds, but result from the backformation or conversion of nominal or adjectival compounds, here *sightseeing, babysitting* and *handmade*.

With regard to morphological structure, **determinative compounds**, in which the first constituent actually modifies the second, stand out as the most common type of compounds. All examples in Table 4.5 above are of this type, which are also known as **endocentric compounds**, which means that the head is actually part of the compound. Counterparts to endocentric compounds are **exocentric compounds**, whose head does not appear in the morphological form of the compound. Although it presumably results from a clipping from *paperback book*, the word *paperback* can be used as an illustration of this type. As we have seen in Section 4.4, the head of *paperback* is located outside the compound, and this is what the term 'exocentric' means. Many exocentric compounds rely on a possessive or 'has a'-relation: consider *redhead* 'person with red hair', *paleface* 'person who has a pale face', *redbreast* 'bird that has a red breast'. Compounds of this type are therefore also known as **possessive compounds** or **bahuvrihi compounds,** a term which comes from the ancient Indian language Sanskrit and exemplifies the phenomenon itself, as it literally expresses the notion of 'having a lot of rice' but means 'rich man'. The third type of compound is also often referred to by a Sanskrit term, **dvandva**, meaning 'pair'. Dvandva compounds are compounds

in which there is no modifier-head relation, but both constituents are considered as heads on a par, e.g. *study-bedroom, singer-songwriter, bitter-sweet* or *deaf-mute*. They are also known as **copulative compounds** when they denote the sum of the two meanings (cf. *bitter-sweet*), or as **appositional compounds** when they combine two different descriptions of the referent (*singer-songwriter*). Unlike determinative compounds, dvandva compounds are typically stressed on both elements of the pair. As is shown in Table 4.6, the three major types of compounds can also be differentiated in terms of the logical relations between the constituents and the compound:

Table 4.6: Survey of types of compounds differentiated by internal morphological and semantic structure

type	internal structure	logical relation (A = first constituent; B = second constituent; AB = compound)	example and paraphrase
determinative compounds	modifier-head structure, endocentric (head is part of compound)	AB is a type of B	*mousemat* 'a mousemat is a type of mat'
bahuvrihi compounds	exocentric (head is found outside compound)	AB is neither A nor B but a type of C	*egghead* 'an egghead is neither a type of egg nor a type of head but a type of person'
dvandva compounds, either copulative or appositive	two-headed structure, endocentric	AB is both A and B	*singer-songwriter* 'a person who is both a singer and a songwriter'

A special challenge for the analysis of morphological structure arises in synthetic compounds of the nominal types *theatre-goer* and *shareholding* and the adjectival types *eye-catching* and *dark-haired*, all of which involve verbal elements and bound lexical morphemes. The problem concerns the branching in binary immediate constituents and the allocation of modifier and head roles. In all four cases, an analysis in terms of a compound consisting of a simple modifier (*theatre, share, eye* and *dark*) and a suffixed head is ruled out, as the potentials heads *goer, holding, catching* and *haired* are at least doubtful with regard to their status as existing lexemes. Analyses in terms of suffixations with complex modifiers (*to theatre-go + -er*, *to sharehold + ing*, *to eye-catch + -ing* and *dark-hair + -ed*) are equally unsatisfactory on the same grounds that the potential bases do not exist. In these cases, and also in those numerable ones where a compound analysis seems at least possible, for instance for *bus driver*, it may seem advisable to argue that compounding and suffixation take place at the same time, so to speak, and to regard these lexemes as synthetic compounds formed by compressing major

components of sentences into one word (cf. *theatre-goer* → 'someone who goes to the theatre').

Further, somewhat less typical classes of compounds include so-called phrase compounds (e.g. *father-in-law, rough-and-ready, man-in-the-street, good-for-nothing*) and particle compounds derived from phrasal verbs (*take-away, breakthrough, handout, take-off*), which present a serious problem for modifier-head analysis. Neoclassical compounds are formations that also combine two concepts in a manner very similar to compounds, but these are not encoded by free lexical morphemes, but rather by bound forms derived from Greek and, less frequently, Latin. Examples of these somewhat learned and often technical words include *democrat, photograph, biography, technology* and *microscope*.

The types of semantic structures and internal relations that can be realized by compounds are virtually unlimited. Nevertheless, some tendencies concerning particularly frequent types can be identified. The examples of root compounds featuring *chair* as head have already given you a glimpse of some of the most dominant relations. More examples are provided in Table 4.7:

Table 4.7: Frequent semantic relations in root compounds

relation	examples
FUNCTION	*gunpowder, breadbasket, toothbrush*
PART-WHOLE	*coat-collar, door-knob, picture-frame*
COMPARISON	*bell skirt, frogman, pot-belly*
TIME	*nightclub, morning coffee, midnight feast*
LOCATION	*water-rat, garden-party, tombstone*
MATERIAL	*ironware, gold ring, stone wall*
CONTAINMENT	*apple cake, sandpaper, picture book*
SOURCE	*spring water, seafood*

Note that the semantic interpretation of compounds, especially root compounds, often offers several equally plausible options.

KEY POINTS: Compounding

- compounds are complex lexemes consisting of at least two free lexical morphemes, i.e. lexemes
- root compounds consist of free morphemes only, while synthetic compounds include bound lexical morphemes

- endocentric compounds include a constituent encoding the head; in contrast, the head is not expressed on the morphological surface of exocentric compounds

- the major types of compounds are determinative, bahuvrihi and dvandva compounds

- compounds can exhibit a variety of internal relations including FUNCTION, PART-WHOLE, COMPARISON and others

Exercise 4.3

Describe the following compounds in terms of their morphological forms and structures and classify them as determinative, possessive or copulative compounds. Provide the information in a table following the two models given:

Compound	Morphological form and structure	Classification
a. *credit card*	*credit card* N N Mod H	determinative noun compound
b. *answering machine*		
c. *lemon-yellow*		
d. *roller blade*		
e. *actor-director*		
f. *hard-working*	*hard work ing* Adj V sfx (Adj) [Mod [Mod H]H]	determinative adjective compound, synthetic
g. *birdbrain*		
h. *shareholder value*		

4.5.2 Prefixation

Prefixation is the word-formation pattern which attaches a bound lexical morpheme at the front of a base, which typically includes at least one free lexical morpheme, i.e. lexeme. Only very few prefixes go hand in hand with a change of word-class: *a-* (*asleep*), *be-* (*beloved*), *en-* (*encourage*) as well as *de-* (*deform*), *dis-* (*displace*) and *un-* (*unsaddle*) in certain uses, while the large majority of prefixes are word-class-maintaining. Prefixation thus has first and foremost semantic, rather than grammatical, effects on a base.

With regard to morphological form and structure, we can distinguish nominal, adjectival and verbal prefixation patterns. Table 4.8 (extracted from Quirk, Greenbaum, Leech & Svartvik 1985: 1540–1546) provides a survey of frequent prefixes, which is organized in terms of semantic groups. It provides information on the word-classes of the bases with which they occur and contains information on their major meanings or semantic relations.

Table 4.8: Frequent types of prefixes (extracted from Quirk et al. 1985: 1540–1546)

semantic type	prefix	meaning	nominal examples	adjectival examples	verbal examples
negative	*a-*	'lacking in'	-	*amoral, asexual*	-
	dis-	'the converse of'	*disorder, discontent*	*disloyal*	*disobey*
	in-	'not', 'the converse of'	-	*incomplete, illogical, irresponsible, impossible*	-
	non-	'not'	*non-smoker*	*non-degradable*	-
	un-	'the converse of'	-	*unfair, unexpected*	-
reversative and privative	*de-*	'reversing the action'	attached to deverbal nouns, e.g. *de-nationalization*	-	*defrost, de-escalate*
		'remove from'	-	-	*delouse, degasify*
	dis-	'reversing the action'	-	-	*disconnect, disinfect*
	un-	'reversing the action'	-	-	*unzip, unpack, unwrap*
		'depriving of'	-	-	*unseat, unmask, unman*
pejorative	*mal-*	'badly', 'bad'	*malnutrition*	*malodorous*	*maltreat*
	mis-	'wrongly', 'astray'	*misconduct*	*misleading*	*mishear*
	pseudo-	'false', 'imitation'	*pseudo-intellectual*	*pseudo-scientific*	-
degree or size	*co-*	'joint'	*co-pilot*	-	*co-exist*
	hyper-	'extreme'	-	*hypersensitive*	-
	mini-	'little'	*mini-skirt*	-	-
	out-	'surpassing'	*outnumber*$_V$	-	*outgrow*
	over-	'excessive'	-	*over-confident*	*overreact*
	sub-	'below'	-	*subnormal*	-
	super	'more than' 'very special'	*superman*	*supernatural*	-
	under-	'too little'	-	*underprivileged*	*underplay*
orientation and attitude	*anti-*	'against'	*anti-war*	*anti-social*	-
	contra-	'opposite'	*contradistinction*	*contrafactual*	*contraindicate*
	counter-	'against'	*counter-espionage*	*counter-clockwise*	*counteract*
	pro-	'for', 'on the side of'	-	*pro-American*	-
locative	*inter-*	'between', 'among'	*inter-war*	*international*	*intermarry*
	sub-	'under'	*subsection*	*subnormal*	*subdivide*
	super-	'above'	*superstructure*	-	*superimpose*
	trans-	'across'	-	*transatlantic*	*transplant*
time and order	*ex-*	'former'	*ex-husband*	-	-
	fore-	'before'	*foreknowledge*	-	*foretell*
	post-	'after'	*post-war*	*post-Freudian*	*postpone*
	pre-	'before'	*pre-war*	*pre-marital*	*pre-heat*
	re-	'again', 'back'	*re-analysis*	-	*rebuild*
number	*bi-, di-*	'two'	*biplane, dioxide*	*bilateral, divalent*	-
	poly-, multi-	'many'	*polytechnic, multiform*	*multi-racial*	-
	semi-, demi-	'half'	*semivowel, demigod*	*semi-conscious*	-
	tri-	'three'	*tricycle*	*tripartite*	-
	uni-, mono-	'one'	*unisex, monoplane*	*unilateral, monosyllabic*	-

Looking at the table, you will perhaps realize that the vast majority of those prefixes that still exist in present-day English and are also still productive and thus used to form new lexemes are of Latin, French and Greek rather than of native Germanic origin. Exceptions are the forms *fore-* as well as *under-, over-* and *out-*. You may also have noticed that the prefix *in-* has a number of variants depending on the first sounds of the base to which it is attached (cf. *indirect, illegal, impossible, irresponsible*). These assimilations often took place in Latin or French before the words were borrowed into English. From a synchronic descriptive point of view, the variants can be considered allomorphs of bound lexical morphemes.

KEY POINTS: Prefixation

- prefixation is a word-formation process in which a bound lexical morpheme is attached to the front of an existing lexeme
- in the vast majority of cases, prefixation does not change the word-class of the base but has an effect on its meaning
- most prefixes in present-day English are not of Germanic origin but come from Greek, Latin and French

---

4.5.3 Suffixation

With regard to morphological form, suffixation appears to be a perfect mirror image of prefixation: a bound lexical morpheme is attached at the end of a base which consists of at least one free lexical morpheme. However, the effects of suffixation on the base are so fundamentally different from the effects of prefixation that to stress this similarity would indeed be rather misleading. Although there are a number of suffixes that keep the word-class of the base intact, including the nominal suffixes *-ship* (*lordship*), *-let* (*droplet*) and *-ing* (*tubing*) and the adjectival suffix *-ish* (*greyish*), these make up a comparatively small portion of the full set of derivational suffixes, whose main function arguably is to bring about a change in word-class. A survey of English suffixes is therefore also more reasonably arranged in terms of their target word-classes, i.e. the word-classes of the products of the derivation process, and the word-class of the base (cf. Table 4.9). The most precise and economical way of describing specific suffixation patterns follows the format '*de*-base target word-class formation'. For example, *signify* would be described as the product of a de-nominal verb-formation, *amendment* as a de-verbal nominalization, *manageable* as a de-verbal adjective formation.

Table 4.9: Frequent types of suffixes (extracted from Quirk et al. 1985: 1546–1558)

noun-forming suffixes	de-nominal	abstract	-age: mileage, footage -ery: drudgery, slavery -ful: spoonful, glassful -hood: brotherhood, widowhood -ing: carpeting, farming -ism: idealism, impressionism -ship: friendship, membership
		concrete	-er: Londoner, villager -ess: actress, lioness -ette: kitchenette -let: booklet, piglet -ster: trickster, gangster
	de-verbal	abstract	-age: drainage, leverage -al: refusal, dismissal -ation: exploitation, exploration -ment: amazement, embodiment
		concrete	-ant: contestant, informant -ing: building, opening -ee: employee, payee -er, -or: driver, writer, computer, actor
	de-adjectival	abstract	-ity: sanity, mobility -ness: happiness, kindness
noun/adjective-forming suffixes	de-nominal or – de-adjectival	reference to persons and membership qualities	-ese: Japanese, Chinese -(i)an: Darwinian, Elizabethan, Russian -ist: violinist, stylist -ite: socialite, Raffaelite
adjective-forming suffixes	de-nominal	native	-ed: wooded, simple-minded -ful: useful, delightful -ish: foolish, snobbish -less: careless, restless -like: childlike, monkeylike -ly: brotherly, friendly -y: sandy, wealthy
		foreign	-(i)al: dialectal, professorial -esque: romanesque, Kafkaesque -ic: atomic, heroic -ous: desirous, ambitious
	de-verbal		-able: washable, debatable -ive: attractive, explosive
adverb-forming suffixes	de-adjectival		-ly: extremely, calmly
	de-nominal		-wards: northwards -wise: clockwise, crosswise
verb-forming suffixes	de-nominal		-ate: orchestrate, hyphenate -ify: codify, beautify -ize: hospitalize, symbolize
	de-adjectival		-en: broaden, harden -ify: simplify, amplify -ize: legalize, publicize

If you study the examples in this list very closely, you will not fail to notice a number of peculiarities about suffixation which deserve special attention: firstly, some suffixes bring about changes in the pronunciation of the base concerning the quality and length of vowels

and/or the allocation of the main stress. Cases in point include *explore – exploration, atom – atomic* and *sane – sanity*. Secondly, a small number of suffixes, mainly *-ee, -ation* and *-esque*, attract the main stress, while others shift it (e.g. *-ic, -ian, -ity*) or leave it unchanged. Thirdly, compared to the large number of noun-forming and adjective-forming suffixes, the list of verb-forming suffixes is quite short. As you will see, this is compensated for by the process of conversion (cf. Section 4.5.4), which has produced massive numbers of verbs derived from nouns and adjectives. Fourthly, unlike prefixes, suffixes frequently occur in sequences of several types, each bringing about a change of word-class. The adjective-forming suffix *-able* is frequently followed by the noun-forming suffix *-ity* (cf. *washability*, *debatability*); the verb-forming suffix *-ize* is added to the adjective-forming suffix *-ar* and frequently followed by the noun-forming suffix *-ation*, cf. *pol(e)-ar-iz(e)-ation*. Finally, from a semantic point of view frequent noun-forming suffixes fall into two basic categories, those producing concrete nouns referring to people and objects (e.g. *-er, -or, -ant, -ee* and *-ing* and those forming abstract nouns (e.g. *-ation, -ment, -age, -ism, -ity* and also *-ing*). Adjectival formations typically refer to qualities and characteristics attributed to people, objects and ideas, and to notions such as ability and potentiality (*-able, -ive*). Verb-forming suffixes show a strong tendency to form transitive verbs incorporating a causative element that can be paraphrased by 'make', cf. *simplify* 'make simple' or *harden* 'make hard'.

As the list in Table 4.9 also suggests, suffixes are of course restricted with regard to the types of bases with which they can combine. In more technical parlance, suffixes – like prefixes in fact – are subject to productivity restrictions (cf. Bauer 2001). These concern first and foremost the word-class properties of bases. For instance, while the suffix *-er* can be added to nouns to form concrete nouns denoting a typical quality of persons or, less frequently, objects (e.g. *Londoner, villager*) and to verbs in order to refer to the agents of actions (*driver, teacher*) or instruments (*computer, dish-washer*), de-adjectival formations (**consistenter, *patienter*) are unacceptable. Knowledge about such restrictions can be very useful in the analysis of word-formation products, because it allows you to predict, for example, that the nominalization *cleaner* must be derived from the verb *to clean* rather than the adjective *clean*. Most suffixes have further productivity restrictions concerning more specific grammatical or semantic properties. The noun-forming suffix *-ee*, for instance, typically combines with bases expressing the patient rather than agent role in a paraphrase: *employee* denotes 'someone who is employed, *interviewee* 'someone who is interviewed'. The adjective-forming suffix *-able*

tends to require transitive rather than intransitive verbs as bases. However, as formations like *sleepable* and *livable* indicate, these productivity restrictions are often not hard and fast rules.

KEY POINTS: Suffixation

- suffixation is a word-formation process which attaches a bound lexical morpheme at the end of an existing lexeme
- in the vast majority of cases suffixation changes the word-class of the base
- suffixation typically creates nouns and adjectives rather than verbs
- suffixes are subject to productivity restrictions

Exercise 4.4

Give descriptions of the following complex lexemes using the format provided in the models:

a. *countless*   de-verbal adjective formation
b. *unfair*   negative adjective prefixation
c. *darkness*
d. *subcategory*
e. *foreigner*
f. *mispronounce*
g. *Australian*
h. *simplify*
i. *carbonize*
j. *re-open*
k. *dishonest*
l. *painting*

4.5.4 Conversion and zero-derivation

Not only can words be transferred from one word-class to another by the addition of a suffix, but this also takes place without any visible changes to their form. The nouns *hammer, bottle* and *father*, for instance, have been turned into verbs with no formal change, and so have the adjectives *clean, tidy* and *dirty*. The process of conversion is made responsible for these changes. It is defined as a word-formation process which transposes a lexeme to a new word-class without the addition of an overtly marked suffix. That this change has actually taken

place can mainly be gleaned from the new grammatical functions that the converted lexeme can fulfil. In many cases, the semantic paraphrase also gives a hint: *to bottle* can be glossed as 'to put into in a bottle', *to father* as 'to act as father to'. These paraphrases can often be used as a guide in determining the direction of derivation of a given conversion, as the base lexeme is usually part of a felicitous paraphrase of the derived lexeme. To paraphrase the noun *father* by something like 'someone involved in an act of fathering' would be decidedly odd; the same applies to a potential paraphrase 'result of an act of cleaning' for the adjective *clean*. Not all products of conversion lend themselves to this test, however. Particularly tricky to work out with regard to the direction of derivation are a huge number of abstract noun-verb pairs including *love, aim, plan, attempt, doubt, hope* and *fear*, which more or less defy all attempts to allocate the roles of base and derivative. The grouping of words of this type in Table 4.10, which gives a survey of the dominant types of conversion (Quirk et al. 1985: 1560–1563), is therefore potentially controversial.

Table 4.10: Frequent patterns of conversion

de-verbal nouns	
'state', 'state of mind'	*desire, dismay, doubt, love, smell, taste, want*
'event/activity'	*attempt, fall, hit, laugh, release, search, swim*
'object of V'	*answer* ('that which answers'), *bet, catch, find*
'subject of V'	*bore* ('s.o./s.th. who/that bores'), *cheat, coach*
'instrument of V'	*cover* ('s.th. with which to cover things'), *paper*
'manner of V-ing'	*walk* ('manner of walking'), *throw*
'place of V'	*divide, retreat, rise, turn*
de-nominal verbs	
'to put in/on N'	*bottle, corner, catalogue, floor, garage*
'to give/provide with N'	*butter (bread), coat, commission, grease, oil*
'to deprive of N'	*core, peel, skin*
'to … with N'	*brake, elbow, fiddle, hand, finger, glue*
'to be/act as N with respect to'	*chaperone, father, nurse, parrot, pilot*
'to make/change … into N'	*cash, cripple, group*
'to send/go by N'	*mail, ship, telegraph; bicycle, boat, canoe*
de-adjectival verbs	
'to make (more) Adj' (trans. V)	*calm, dirty, dry, humble, lower*
'to become Adj'(intrans. V)	*dry, empty, narrow, weary (of), yellow*

There has been some controversy in linguistics as to how the phenomenon of conversion should be explained theoretically. The approach presented so far regards it as some kind of invisible transfer process very similar to overt suffixations. Another possibility is to argue that many lexemes in the English lexicon have either multiple word-class membership or are not determined with regard to word-classes anyway, but only acquire word-class properties when

used in syntactic contexts (Farell 2001). Extreme cases supporting the latter view include the form *round*, which can be used as a noun, verb, adjective, adverb and preposition or particle. Yet another approach, which emphasizes the similarity to overt suffixation, works with the notion of **zero-morpheme** and claims that the word-class change is brought about by a suffix that does not have a formal substance (Marchand 1969: 360ff). The verb *to empty*, for example, would be explained as a result of the addition of the zero-morpheme to the adjective *empty* (*empty – empty + {Ø}*) in analogy to cases like *legal – legalize* or *public – publicize*. An argument in favour of this approach is that the semantic change effected by the zero-morpheme ('cause to be empty') is also identical to the one taking place in overt formations.

KEY POINTS: Conversion and zero-derivation
- conversion is a word-formation process which transfers a lexeme to a new word-class without the addition of an overtly marked suffix
- while conversion frequently produces verbs from nominal or adjectival bases, for many cases it can be difficult to determine the direction of derivation
- an alternative account is the idea of zero-derivation which claims that a zero-morpheme is responsible for the observable change of word-class

---

Exercise 4.5

Paraphrase the meanings of the following conversions as illustrated in a) and b):
  a. *pocket* V ← *pocket* N: 'put in N'
  b. *kick* N ← *kick* V: 'an act/instance of V-ing'
  c. *nail* V ← *nail* N
  d. *progressive* N ←*progressive* Adj
  e. *grant* N ← *grant* V
  f. *model* V ← *model* N
  g. *warm* V ← *warm* Adj

4.5.5 Back-formation

Like suffixation and conversion, back-formation is a word-class-changing process. In contrast, however, this process is not marked by the addition of morphological material or by keeping the surface form unchanged, but by the elision of material. It can be defined as the process whereby the deletion of a morpheme or morpheme-like element results in a

transposition of a lexeme to a new word-class. The verb *babysit*, which is derived from the noun *babysitter* by means of back-formation, and the verb *to sightsee* derived from *sightseeing* have already been mentioned above. Other frequently quoted examples include *to burgle → burglar*, *to edit → editor, to laze → lazy* and *to televize → television*. As these examples indicate, back-formation joins conversion as a predominantly verb-forming process.

Of course, back-formation is much more difficult to spot and identify than additive word-formation processes. Deficient inflectional paradigms can be an indicator, as is indicated by the rather unusual forms *I babysat last week* or *we sightsaw in London*. Paraphrases can also contribute as an argument: while 'to sit by the baby' is a rather poor gloss of *to babysit*, a paraphrase that includes the noun *babysitter*, e.g. 'to act as a babysitter' is quite plausible. If you keep in mind that genuine verbal compounds do not exist in English, this will also help you to realize that verbs like *bottle-feed, house-hunt* or *chain-smoke* might be the results of a back-formation process.

4.5.6 Non-morphemic word-formation types

The hallmark of non-morphemic word-formation processes, in addition to the fact that they do not obey morpheme boundaries, is that they are less regular and therefore less predictable. This means that given a recent verb like *to desktop-edit*, you can easily envisage that someone may eventually find it convenient to coin the adjective *desktop-editable*. However, knowledge of the words *floor* and *wardrobe* will not have put you in a position to foretell that someone has actually found it funny to coin the blend *floordrobe* to refer to an untidy room where lots of clothes are scattered all over the floor.

Four main types of non-morphemic word-formation process are commonly distinguished: blending, clipping, acronym-formation and reduplication.

The term blending subsumes a number of ways in which two or more words can be merged or telescoped into each other. In the most typical cases, overlapping segments of words are exploited, as for example in the classic *smog → smoke* and *fog* or the more recent *wintertainment → winter* and *entertainment*, and the meanings of the blended lexemes are also blends of the meanings of the source lexemes. In *sexploitation*, both source words are retained in their full forms. The example *floordrobe* mentioned above illustrates the type of blend in which there is no or only a very superficial kind of overlap. A further example is the

well-known noun *brunch* →*breakfast* and *lunch*, which differs from *floordrobe*, however, in that *floor* finds its way into the blend without being subject to a change. From a formal perspective, the latter example could in principle also be seen as a combination of *floor* and *drobe*, a clipped form of *wardrobe*, but the merged meaning of 'floor that serves as a wardrobe' speaks against that.

The process of clipping is responsible for a number of entirely common everyday words, whose sources are hardly known today. For example, *car* is a front clipping of *motor car* and *bus* a front clipping of *omnibus*; *pub* is a back clipping of *public house* and *zoo* a back clipping of *zoological garden. Flu* has emerged from *influenza* by way of a combination of front and back clipping.

Acronym-formation is an extremely productive process, especially in technical and institutional registers, but also increasingly in youth language and computer-mediated communication (cf., e.g., *FAQs* ← *frequently asked questions*, *lol* ←*laughing out loud, brb* ← *be right back* and many other examples). Regarding the pronunciation of these formations we can distinguish those that are pronounced as words, e.g. *NATO* (← *North Atlantic Treaty Organization), AIDS* (←*Acquired Immunodeficiency Syndrome), PEN* (← *poets, essayists, novelists*), from cases where the letters are pronounced separately (e.g. *TV* ← *television*, *UK* ← *United Kingdom*, *BBC* ← *British Broadcasting Corporation*). The former are sometimes labelled as acronyms in a narrow sense, the latter as **initialisms** (cf. Bauer 1983: 223). Usually, the capital letters are used as a sign that a compound or phrase has been reduced to the initial letters, but there are also highly lexicalized forms like *radar* (from **r**adio **d**etection **a**nd **r**anging) or *laser* (from **l**ight **a**mplification by **s**timulated **e**mission of **r**adiation).

Finally, **reduplication** is a quite minor type of word-formation pattern illustrated by lexemes such as *hush-hush, hip-hop* and *walkie-talkie*. As the examples indicate, the pattern subsumes cases where an element is repeated in identical form (*hush-hush*), cases where we have a vowel change (*hip-hop*) and those where the two components rhyme (*walkie-talkie*).

> **KEY POINTS**: Back-formation and non-morphemic word-formation types
>   - back-formation is a word-class changing word-formation process which deletes a morpheme or morpheme-like element
>   - in blending, the forms and meanings of words are merged
>   - in clippings, parts of words are deleted without a change in meaning
>   - acronyms and initialisms are shortened forms retaining the initial letters of compounds and other fixed sequences of words; the former are pronounced as words, the latters as sequences of letters
>   - reduplication is a fairly rare word-formation process repeating a word or word-like element either identically or in a slightly varied form

Exercise 4.6

Classify the following lexemes in terms of their formation pattern:

   a. *tick-tick*
   b. *ad*
   c. *Oxbridge*
   d. *USA*
   e. *lab*
   f. *higgledy-piggledy*
   g. *grannie*
   h. *IRC*
   i. *prefab*
   j. *fanzine*
   k. *fridge*
   l. *IMO*
   m. *hi-fi*

4.6 A note on theoretical issues

Although it may not have struck you while reading the chapter, the approach presented here is largely a practical, down-to-earth one, which focuses on the methods and background knowledge required to carry out morphological analyses and appreciate the system behind the

structures of words. Only very little has been said about the manifold theoretical disputes concerning the precise characteristics of this system. Back in the 1970s and 1980s, the discussion was dominated by the question of whether morphology and word-formation work essentially on the basis of principles similar to those postulated for syntax, thus producing members of phrasal categories, or whether morphology and word-formation have their basis in the lexicon, the storehouse of lexical categories. A more recent controversy concerns the format of the system described in this chapter and, specifically, the nature of the morphological knowledge which individual speakers and speech communities as a whole apparently have at their disposal – otherwise they would constantly coin ill-formed words. For a long time, this knowledge was modelled in the form of strict and abstract rules operating over entities defined in terms of equally abstract categories; work on morphology was very much preoccupied with defining these rules and the prerequisites for their input, and with determining the nature of their output. Individual words and how they are coined, used and propagated had hardly any role to play in this model. More recently, this approach has been rivalled by one which proceeds from the assumption that morphological knowledge is available in the form of more flexible schemas (Bybee 2007; Kemmer 2003) or constructions (Booij 2010) which are extracted or distilled by speakers from their constant exposure to inflected word-forms and complex lexemes (Schmid 2011: 85, 93–95). While these schemas provide them with the knowledge to distinguish well-formed from ill-formed novel creations, speakers are still free to coin creative new words but will then be more likely to have to face the possibility that their creations are not taken up by other speakers and therefore do not catch on.

4.7 Summary

This chapter has introduced you to inflectional morphology and word-formation, including derivational morphology. You should now be familiar with different types of morphemes and their characteristics, as well as different types of word-formation patterns.

The section on inflectional morphology has shown that languages differ considerably with regard to the extent to which they rely on inflectional morphemes to mark grammatical categories and relations. Different examples that will help you remember the major types were English (increasingly analytic), German (more synthetic than English), Latin (fusional) and Turkish (agglutinating). Complexity in the field of inflectional morphology is not only caused by the number of inflectional morphemes but also by their variants, the allomorphs.

The section on word-formation has demonstrated the need to analyze complexe lexemes systematically at the levels of morphological form, morphological structure and semantic structure. You have seen that the more regular, morphemic word-formation types of compounding, prefixation, suffixation and conversion differ in their effects on the grammatical, formal and semantic structures of the elements involved. Essentially, the function of compounding is to join words and concepts in order to create more specific words and concepts; the main function of prefixation is to modify the meanings of existing words in a number of basic ways; the main effect of suffixation and conversion is to change the word-class of the base. The non-morphemic word-formation patterns (blending, clipping, acronym-formation and reduplication) are less regular and more creative, sometimes also more playful, than the morphemic ones. The idea, hinted at in the final section, that knowledge about word-formation patterns is available in the form of rough schemas and blueprints rather than strict rules seems particularly plausible for these flexible formation types but is also applicable to the more regular ones.

Suggestions for further reading

Bauer, L. (2003). *Introducing linguistic morphology*. 2nd ed., Edinburgh: Edinburgh University Press.
An accessible, richly illustrated introduction to linguistic morphology containing examples from a wide range of European and other languages. The book covers inflectional and derivational morphology as well as compounding and discusses the relation of morphology to phonology.

Bauer, L., Lieber, R. & Plag, I. (2013). *English morphology: A reference guide to contemporary English word-formation and inflection,* Oxford: Oxford University Press.
This handbook, written by three of the leading authorities in the field of English morphology, is a very useful resource for finding out more about all aspects relating to the field of morphology.

Schmid, H.-J. (2011). *English morphology and word-formation. An Introduction*, Berlin: Erich Schmidt.
This introduction covers inflectional morphology from a synchronic and diachronic point of view, as well as word-formation. The discussion is based on authentic examples taken from a

corpus, which is also used for quantitative analyses of data. With regard to theory, special emphasis lies on cognitive-linguistic approaches and socio-pragmatic aspects, while the generative tradition is clearly kept in the background. One chapter offers a systematic description of the processes involved in the establishment of new words.

References

Bauer, L. (1983). *English Word-formation*, Cambridge: Cambridge University Press.

Bauer, L. (2001). *Morphological Productivity*, Cambridge: Cambridge University Press.

Bauer, L. (2003). *Introducing linguistic morphology*. 2nd ed., Edinburgh: Edinburgh University Press.

Bauer, L., Lieber, R. & Plag, I. (in press). *English morphology: A reference guide to contemporary English word-formation and inflection,* Oxford: Oxford University Press.

Booij, G. (2010). *Construction Morphology*, Oxford: Oxford University Press.

Bybee, J. (2007). *Frequency of Use and the Organization of Language*, Oxford: Oxford University Press.

Farell, P. (2001). Functional shift as category underspecification. *English Language and Linguistics, 5*, 109–130.

Kemmer, S. (2003). Schemas and lexical blends. In: H. Cuyckens, Th. Berg, R. Dirven and K. –U. Panther (Eds.), *Motivation in language: Studies in honor of Günter Radden* (pp. 69–97). Amsterdam: Benjamins.

Lieber, R. (2005). English word-formation processes. In P. Štekauer & R. Lieber (Eds.), *Handbook of word-formation* (pp. 375–427). Dordrecht: Springer.

Marchand, H. (1969). *The categories and types of present-day English word-formation. A synchronic-diachronic approach* (2nd ed.), Munich: Beck.

Plag, I. (2003). *Word-formation in English*, Cambridge: Cambridge University Press.

Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A comprehensive grammar of the English language*, London – New York: Longman.

Schmid, H.-J. (2011). *English morphology and word-formation. An Introduction*, Berlin: Erich Schmidt.

Stockwell, R. & D. Minkowa (2001). English Words – History and Structure. Cambridge: Cambridge University Press.

Glossary

acronym-formation: non-morphemic word-formation process forming new words by reducing compounds and phrases to their initials; **acronyms**, in the narrow sense, are products of this process which can be pronounced like normal words, e.g. *NATO, AIDS.*

agglutinating language: type of synthetic language (e.g. Turkish) which is rich in inflectional morphemes and produces long words by concatenating sequences of morphemes that do not overlap and typically have a one-to-one relation of form and meaning.

allomorph: different formal realizations of morphemes, especially inflectional ones.

analytic language: type of language which does not encode grammatical categories by means of inflectional morphemes affixed to stems but by other means such as word order, auxiliaries and particles.

appositional compound: a type of two-headed, non-determinative compound combining two different descriptions of the same referent (e.g. *singer-songwriter*).

back-formation: typically word-class-changing word-formation process involving the deletion of a suffix or suffix-like element (e.g. *to babysit ← babysitter*).

bahuvrihi compound: type of exocentric compound where the meaning of the compound stands for a property or part of the head (e.g. *paleface* 'person who has a pale face').

blending: type of non-morphemic word-formation process involving the merging or telescoping of two lexemes into one.

bound morpheme: type of morpheme that cannot occur by itself but only as an affix to a free morpheme.

clipping: type of non-morphemic word-class-preserving word-formation process involving the deletion of parts of the source lexeme at the end (back-clipping, e.g. *photo ←*

photograph), front (front-clipping, e.g. *bus ← omnibus)* or on either side (e.g. *flu ←*
influenza).

complex lexeme: lexeme consisting of more than one lexical morpheme.

compounding: type of morphemic word-formation process involving the combination of at
least two free lexical morphemes, i.e. lexemes.

conversion: type of typically word-class-changing word-formation process which transposes
a lexeme to a new word-class without the addition of an overtly marked suffix (e.g. *empty* V
← *empty* Adj).

copulative compound: a type of two-headed, non-determinative compound denoting the sum
of two meanings (e.g. *bitter-sweet*).

derivational morpheme: type of morpheme involved in the creation of new lexemes,
typically close to the stem and subject to productivity restrictions.

derivational morphology: branch of morphology dealing with word-formation types using
prefixes and suffixes.

determinative compound: type of endocentric compound exhibiting a modifier-head relation
between the constituents.

dvandva compound: type of compound which has two heads (e.g. *bitter-sweet, actor-*
director).

endocentric compound: type of compound where one constituent, in English usually the
final one, encodes the grammatical and semantic head.

exocentric compound: type of compound whose grammatical and semantic head is encoded
by neither of the constituents but 'lies outside' the compound (e.g. *redbreast*).

free morpheme: type of morpheme that can occur as a free form.

fusional language: type of language (e.g. Latin) which is rich in inflectional morphemes frequently encoding several meanings in one form.

grammatical morpheme: type of morpheme encoding grammatical meanings and relations.

head: constituent of a complex lexeme which determines its grammatical and fundamental semantic properties.

inflectional morpheme: type of bound morpheme creating word-forms and marking grammatical categories and relation, typically positioned at the very end of words.

inflectional morphology: branch of morphology dealing with the bound morphological markers of grammatical categories and relations.

initialism: product of acronym-formation which is pronounced as a sequence of individual letters (e.g. *IRC* ← *internet relay chat*).

lexical morpheme: type of morphemes encoding rich conceptual meanings.

modifier: constituent of a complex lexeme which specifies the head.

morpheme: smallest meaning-bearing unit of a language.

morphological conditioning: dependence of the choice of allomorphs on the final morpheme of the stem; pertains to 'irregular' allomorphs, e.g. *sang* or *mice*.

phonological conditioning: dependence of the choice of allomorphs on the final phoneme of the stem; pertains to 'regular allomorphs, e.g. *kissed* or *dogs*.

possessive compound: type of exocentric compound based on a possessive relation between the meaning encoded by the compound constituents and the meaning of the exocentric head (e.g. *redbreast* ' a bird that has a red breast').

prefixation: type of typically word-class-preserving word-formation process involving the attachment of a bound lexical morpheme at the front of a base (e.g. *unfair, disagree*).

reduplication: non-morphemic word-formation process involving the repetition of a word or word-like element in unchanged form (e.g. *hush-hush*), with a different vowel (e.g. *hip-hop*) or a different consonant (e.g. *boogie-woogie*).

root compounds: type of compound consisting of free lexical morphemes only, as opposed to synthetic compounds (e.g. *doorknob, lamppost*).

simple lexeme: lexeme consisting of one lexical morpheme only.

suffixation: type of typically word-class-changing word-formation process involving the attachment of a bound lexical morpheme at the end of a base (e.g. *fairness, agreement*).

synthetic compound: type of compound consisting of at least three morphemes, among them one bound morpheme and one encoding a verb (e.g. *washing-machine*, *dog owner*); many synthetic compounds cause problems for analyses based on binary branching.

synthetic language: type of language which encodes grammatical categories by means of inflectional morphemes.

word-formation: cover term for morphemic and non-morphemic processes involved in the creation of new lexemes on the basis of existing morphemes and lexemes; branch of morphology dealing with these processes.

zero-morpheme: theoretical construct used, among other things, to explain conversion as a form of derivation comparable to (overt) suffixation, e.g. *empty* V $\leftarrow$ *empty* Adj + {Ø}.