1 Introduction

1.1 What Is Word-Formation?

'Word-formation' is the standard label for the way in which words are built up from smaller recurrent formal elements. These smaller formal elements are generally presumed to be linked in a fairly direct way to the meaning of the word that is formed from them. Word-formation is widely assumed to be made up of at least two distinct types: compounding or composition, the way in which compounds like *molehill* and *threadbare* are constructed from smaller words; and derivation, the way in which derivatives like *unfriendly* and *discovery* are constructed from prefixes, suffixes and a word which is their base.

Given the well-known ambiguity of the term 'word' in modern linguistics, it might be better if word-formation were termed 'lexeme-formation'. However, the term 'word-formation' was established before the term 'lexeme' became normal, so 'word-formation' is the usual term. But the term 'word-formation' is also sometimes used to include inflectional morphology, and so to deal with the construction of word-forms such as *covered* and *elephants*; this means that precisely what is included under word-formation is not necessarily fixed (see Chapter 2). The standard notation is that *covered* (in italics) is a word-form belonging to the paradigm of the lexeme COVER (in small capital letters), but this depends on several assumptions being agreed to (see Section 2.4).

Some people prefer the term 'lexical morphology' to 'word-formation' (e.g. Coates 1987). This seems to imply that there is nothing in word-formation that is not part of morphology. The difficulty with this is that there is not necessarily agreement on what the term 'morphology' encompasses, either. While everyone agrees that compounding, inflection and derivation are part of morphology, it is not necessarily true that everyone agrees that the formation of words by conversion (e.g. the link between the verb *to cuddle* and the noun *a cuddle*) is part of morphology, and neither is necessarily true that everyone agrees that the formation of initialisms such as *MIT* from *Massachusetts Institute of Technology* (note that the initialism is not *MIOT*) is part of morphology. Yet both of these might well be included as part of word-formation.

1

We can already see that word-formation is plagued by terminological prob-
lems, and some of these will be expressly addressed in this book (see e.g.
Chapters 9, 20 and 22). Terminological problems cannot be solved by fiat. But
if they are made explicit, authors can position themselves in regard to the major
points of contention involved.

Word-formation can be viewed from at least two different viewpoints. It can
be viewed as dealing with the analysis of known words, or it can be viewed as
dealing with the ways in which new words are created by speakers. Since the
1970s – possibly before – the predominant approach has been the latter, at least
overtly. However, most handbooks provide many examples of the former, and
the two viewpoints are not necessarily kept strictly apart. Both approaches will
be considered here in different chapters. The latter perspective gives rise to the
question of productivity, whether and to what extent individual morphological
processes (including compounding and the use of prefixes and suffixes) are
used to create previously unknown words. Productivity has proved to be a very
difficult area of study (see Bauer 2001). Despite arguments made, for instance
by Bauer (1983), productivity is often seen as being one of the ways in which
the creation of new words differs from the creation of new sentences. But just
how variable productivity works, how it can be measured (if at all) and how it
may be constrained are still questions that can be debated, and some of these
questions will be raised below (e.g. in Chapter 5).

1.2   Is English Word-Formation Different from Word-Formation in Other Languages?

The short answer to the question raised in the heading here is 'yes'. All languages
differ in the details of how their word-formation works, what semantic categories
are marked in the word-formation (Japanese and Swahili mark causatives overtly
in their morphology, while English tends not to), what kinds of formal means are
used in the creation of words (English tends not to use infixation or reduplicating
prefixes while many other languages do), and what kinds of pattern are found
frequently and what kinds are found rarely. But that is not the reason for having
a book about English word-formation. In principle, I could have written about
word-formation in a different language, or across languages, and either of those
topics might have made a useful contribution. But English is not only a familiar
language for many linguists for whom it is not a first language, it is a language for
which a great deal of data is available, in the form of dictionaries (most notably
the *Oxford English Dictionary*), corpora illustrating usage, wordlists, coverage of
word-formation specifically and easily available examples of real usage by
language users. It is also the language for which I have the best intuitions
(however dangerous intuitions on usage may be). For all these reasons, focusing
on English provides materials which allow me to do what I want to do in this

book, namely examine the ways in which word-formation can be described and the problems associated with them. This also means, I hope, that many of the questions that are raised here with specific reference to English will actually be of relevance to other languages and allow for discussions on the ways in which languages can differ. Sometimes, for clarity, comparisons are made with other languages (as has already been done above, see, in particular, Chapter 26), but the focus is on the way in which these shed light on what happens in English.

1.3   The Historical Context

Any book is written in a context that is constantly changing. In particular that context involves the way in which scholars understand the topic that they are dealing with, involving the theories they believe in, the terminology they use and the elements that they perceive as being relevant. The historical context, however, is not necessarily uniform across a field such as linguistics. Ideas which are no longer viewed as current in one area of linguistics may still be considered standard in another. This is inevitable in any subject in which the theories that were dominant when one generation of scholars was trained are no longer dominant when new scholars are trained one or two generations later. In the case of word-formation, the perceived importance of morphology in a grammar has changed markedly over just a few generations, going from being a central aspect of grammar to a marginalized area of study and back to a central topic of focus within linguistics (Coates 1987). There is little point in going through the changes in underlying philosophy which have led to this position, but to consider the different structures and fundamental notions that have been left behind might clarify some of the ways in which word-formation has been dealt with.

The twentieth century provides a textbook example of the way in which linguistic entities can come into and go out of fashion, when we consider the notion of the morpheme. The morpheme seems to have been named in the late nineteenth century, although the notion had been available before then. There were two views of the morpheme, a European one and a North American one, but they merged on the North American pattern. A morpheme is a meaningful element of a word. If we consider the word *unfriendly*, we can see that the word is made up of three elements. The core of the word is *friend*, *friend* is turned into an adjective by the addition of the suffix *-ly*, and *friendly* is made negative by the addition of the prefix *un-*. Each of *friend*, *-ly* and *un-* either is a morpheme or, in slightly later analyses, represents a morpheme. If we follow the latter view, a morpheme is an abstract unit (not a form which can be heard or written), realized by one or more forms. If we follow the European tradition (following Saussure 1916), the morpheme is a minimal sign: it has a form and it has a meaning, but contains nothing smaller with the same qualities. If we follow

the North American tradition (following Bloomfield 1935), a morpheme is a minimal unit of grammatical analysis; again, it contains nothing with the same qualities. In both cases, the notion of the morpheme was modelled on the notion of the phoneme. In the European tradition, it might have bound variants, in the North American tradition it can be viewed as having allomorphs. The bound variants or allomorphs are variant forms which are restricted to particular environments. Consider the words *embark* and *entrain* meaning, respectively, 'to put onto a ship' and 'to put onto a train'. Each is made up of two forms, *em-* and *bark*, *en-* and *train*. The difference between *em-* and *en-* has nothing to do with the meaning, but the bilabial nasal appears before a bilabial consonant, while the alveolar nasal (at least in the written form) is the default case, occurring where there is no need for *em-*. Because of their distribution, the two do not contrast and they also share a meaning, and so can be seen as variants of the same abstract item. They are distinct morphs representing the same morpheme. (Actually, I have simplified a little, because we find *enmesh* more frequently than we find *emmesh*, because /nm/ is a permissible medial consonant sequence – e.g. in *enmity* – while /nb/ would not be possible in **enbark*.) By the 1950s, the morpheme was a standard element of grammatical description, and worked brilliantly for a great deal of English word-formation (and morphology more generally in many languages), exploiting the very strong notion of allomorphy, very much in the structuralist tradition.

Scholars of the time knew that there were problems with this picture, but felt that the notion was valuable enough for it to be worthwhile working round the difficulties these caused. However, more and more problems were found, and as scholars wanted to write explicit grammars in the innovative Chomskyan tradition that became dominant in the 1960s, the problems became viewed as insuperable. This happened first in the description of inflectional systems, although similar examples from derivational morphology could be used to make the same points. The problems were raised and elaborated in works such as the hugely influential Matthews (1972) and Anderson (1992), and can be found summarized in works such as Anderson (2015) and Bauer (2016, 2019). The result was that by the 1990s, the morpheme as it had been presented by the structuralists was no longer considered to be a tenable theoretical notion, at least within theoretical morphological studies, though psycholinguists continued to work with the notion.

Although this is not the place to rehearse all the arguments against the morpheme, we can say briefly that we find instances where a single morph is associated with more than just one meaning, instances where a single meaning is associated with several morphs; we find instances where we have a recognizable morph, but the meaning usually associated with it is not associated with a particular word in which it occurs; we find morphs which do not appear to carry any meaning at all and meanings which do not appear to have

any morph for them to be associated with; we have instances where a meaning appears to be associated with a process rather than with a form, and many instances where it is not clear how far we can stretch the notion of allomorphy. This cumulation of problems leads to difficulties in applying the notion of morpheme to many languages, but then the question becomes what can we replace the morpheme with.

In some instances, particularly in cases of complex inflectional systems, it seems to be possible to build up some morphs by the application of phonological rules (Matthews 1972). Meanings can be associated with whole words or sequences of words rather than with individual morphs (Matthews 1972, Booij 2010). In other cases, we might observe formal patterns which do not directly correlate with particular meanings, but which nevertheless tend to be meaning-bearing (Aronoff 1994) or, if not, at least important in morphological patterning and sometimes more diachronically stable than the individual words that instantiate them. Note in all this that while morphemes may not be used, morphs – the formal side of morphemes – tend to persist, and what is lost is the direct link between individual forms and individual meanings.

Ironically, against such a background, the study of word-formation in English, with its concentration of prefixes and suffixes, seems stuck on the idea of the morpheme. There is still a view that words are made up of formal elements, to each of which can be attributed a meaning or a function. This looks remarkably like analysis into morphemes except that there is recognition that not all morphs have meaning (the -t- in *dramatist* is either a meaningless extender or is part of an allomorph of *drama* which occurs before certain suffixes), that not all morphs have a consistent meaning (the -er in *dishwasher* can denote either a person or an instrument, although some less specific meaning may allow for a single gloss to cover both), we can have meaning spread over multiple morphs (as in *enliven* where prefix and suffix together provide the causative meaning) and more than one meaning in a single morph (as in *song*, which can be seen as containing the meaning of 'sing' as well as the meaning of 'noun') and we can have meanings not associated with any morph (as with the difference between the verb to *whisk* and the noun a *whisk*). That is, those who study word-formation seem more willing than those who study inflection to work round the problems associated with the morpheme rather than simply discarding the notion. Discussions which seem to assume a morpheme-like unit will also be found in this book.

Developments in semantic theory have also had a strong influence on the study of word-formation. The notion of prototype (Rosch 1973, Taylor 2003) has influenced the way in which we view not only the meanings of words, but also the way in which we envisage categories. For instance, we can now think of the suffix -er as having a prototypical meaning centring on the notion of agency (*discoverer*), but fading off into instruments (*sharpener*), experiencers

(*lover*), locations (*diner*) and even patients (*keeper*) rather than as having a set of fundamentally unrelated meanings. Where categories are concerned, we can think of a class such as compounds having more central members and less central members, rather than being defined by a fixed set of criteria or tests. If *girl Friday* is a compound, it is not as central a compound as *flower-girl*, because *girl Friday* is left-headed (see further Chapter 13). An alternative is to see a category such as compound as a canonical category, again with more and less central members of the category (Corbett 2010). Within Cognitive Linguistics, the importance of figures of speech has been seen as central for the semantics of word-formation because so many words involve figures of speech. For instance, a *whirlybird* 'helicopter' is not literally a bird, but resembles a bird in that it can fly, a *jailbird* 'prisoner' is not literally a bird, but is like a budgerigar in that he or she is kept in a cage, and a *thunderbird* is not a bird because it is a mythical creature, once believed to cause thunder. Recognizing the figurative expressions for what they are makes it easier to explain the often complex meanings of words created by word-formation, as well as the meanings of simplex words.

Corresponding to these various developments, there have been many different theoretical approaches to morphological study. In the early 1960s, within Chomskyan generative grammar, morphology was seen as part of syntax, in line with the view of morphemes that was still current at the time. Lees (1960) provides an illustration of this general approach. Later (most obviously in Chomsky and Halle 1968, but much discussed before that date), some of the workings of morphology were subsumed in phonological rules, in a movement that eventually led to Lexical Phonology and Morphology. Viewing morphology as syntactic persists in Distributed Morphology (Halle and Marantz 1994 and a large amount of more recent work). This contrasts with a movement towards seeing morphology as an independent area of linguistics, not just part of other larger fields, a view which is made overt in the title of Aronoff (1994), but is also to be seen in a lot of works where the notion of morphological paradigm is given a central position, both in inflection and in derivation. This point of view is found in various theoretical positions starting with Bybee (1985) and is increasingly important today.

While this brief outline does not cover all the theoretical positions that have been taken over the last century or so (for more detailed coverage, see Stewart 2015), readers can expect to find individual scholars taking positions which call on several of these approaches in slightly different ways, and giving focus to different aspects of the ways in which word-formation can be dealt with. There is no consensus, but there is much to be learned from studies which take different perspectives on the topic.

1.4   Why Study Word-Formation in Particular?

We could ask why we study anything at all, from how earthworms mate to why the universe is expanding, and I imagine that the answer would be much the same: because it's fascinating and we are curious about how the universe works. One of the topics within the range that we could wonder about is why we should study language. Again, its fascination could be part of the answer, but we might have some extra reasons (just as there might be extra reasons for the study of anything). In the case of language, we could argue that one of the strangest things about human beings is not only that they communicate by means of language, but that it is hard to stop them talking! While other species undoubtedly communicate, the human communication system is far more complex than any other we understand, and it is definitional of human beings. We might be classified as *Homo loquens*. Studying language is therefore one way to try to understand part of what it is that makes us human. More than that, most of us speak more than one language, and those people who do not are at least aware that although language may be a part of what defines humankind, it is not always the same language, and that various languages may well be totally incomprehensible to people who have at least one language to help them communicate with others. Most of us, however many or how few languages we speak, will also be aware that if we cannot use language to communicate, our ability to communicate at all drops drastically. Such observations raise many questions, including: what do languages share? Are all languages really just dialects of Human? Or are they so radically different that knowing one is little help in learning another? Is it always possible to translate between languages? Can we learn Dog and can dogs understand Human (see Anderson 2004)?

Once we have decided to study language, we still have a huge problem: languages have so many properties that it is hard to know where to start studying them. Do we treat languages as organized sound (gesture in the case of sign languages), as words chained together, as a means of transferring meaning by having different ways of expressing a huge number of meanings, as a way of organizing human interaction in such a way as to promote a coherent society, as a way of reflecting the societal structures within which we operate, as some kind of code that keeps changing? Language has all those aspects. Reducing our focus to word-formation has some benefits as a way into this morass of complex interactions: it concerns words, which speakers of European languages at least think they have some understanding of; it deals not only with forms but also with meanings; although we can invent new words, most messages are made up of familiar ones, while most sentences (at least in academic discussions) are not at all familiar; and last but not least, words are fun. Many a comedian has asked questions like if a vegetarian eats vegetables,

what does a humanitarian do, or if *sink*, *sank* and *sunk* are all related words, are *think*, *thank* and *thunk* related in the same way? Language is endlessly fascinating, and word-formation, quite apart from being interesting in its own right, makes a good entry point into the wider field of study.

1.5   This Book

Although this book is about word-formation, specifically about word-formation in English, it is not a textbook in the sense that it does not attempt to provide a systematic and thorough discussion of the topic. The title *Reflections* is intended to signify that it covers topics that I have found to be of particular interest in the fifty years I have been thinking and writing about the area, but also that the aim is to provide new insights or points of discussion, not just to provide a summary of the state of the art. The topics covered are loosely collected into thematic groupings, but are quasi-independent: the chapters can be read in isolation, although places where the individual chapters are linked are indicated, as they already have been above. Some matters which may appear minor are covered, some which may seem more important are not. But while topics which are vital to various theories about the way word-formation might work are covered, explanation of individual theories are in general not covered: theories come and go, but how word-formation is actually used lies at the heart of the theories.

Challenge

Find examples from English derivational morphology where the morpheme does not function in the way that is expected, as set out in this chapter. How important are these examples in the general scope of word-formation?

References

Anderson, Stephen R. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.
  (2004). *Dr Dolittle's Delusion: Animals and the Uniqueness of Human Communication*. New Haven, CT: Yale University Press.
  (2015). The morpheme: Its nature and use. In Matthew Baermann (ed.), *The Oxford Handbook of Inflection*. Oxford: Oxford University Press, 11–33.
Aronoff, Mark. (1994). *Morphology by Itself*. Cambridge, MA: MIT Press.
Bauer, Laurie. (1983). *English Word-Formation*. Cambridge: Cambridge University Press.
  (2001). *Morphological Productivity*. Cambridge: Cambridge University Press.

(2016). Classical morphemics: Assumptions, extensions, and alternatives. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge Handbook of Morphology*. Cambridge: Cambridge University Press, 331–55.

(2019). *Rethinking Morphology*. Edinburgh: Edinburgh University Press.

Bloomfield, Leonard. (1935). *Language*. London: Allen & Unwin.

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

Bybee, Joan. (1985). *Morphology*. Amsterdam: Benjamins.

Chomsky, Noam & Morris Halle. (1968). *The Sound Pattern of English*. New York: Harper & Row.

Coates, Richard. (1987). Lexical morphology. In John Lyons, Richard Coates, Margaret Deuchar & Gerald Gazdar (eds.), *New Horizons in Linguistics 2*. London: Pelican, 103–21.

Corbett, Greville G. (2010). Canonical derivational morphology. *Word Structure* 3, 141–55.

Halle, Morris & Alex Marantz. (1994). Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.), *The View from Building 20*. Cambridge, MA: MIT Press, 111–76.

Lees, Robert B. (1960). *The Grammar of English Nominalizations*. Bloomington: Indiana University Press.

Matthews, P.H. (1972). *Inflectional Morphology*. Cambridge: Cambridge University Press.

Rosch, Eleanor H. (1973). On the internal structure of perceptual and semantic categories. In T.E. Moore (ed.), *Cognitive Development and the Acquisition of Language*, New York: Academic Press, 111–44.

Saussure, Ferdinand de. (1916). *Cours de linguistique générale*. Paris: Payot.

Stewart, Thomas W. (2015). *Contemporary Morphological Theories*. Edinburgh: Edinburgh University Press.

Taylor, John. (2003). *Linguistic Categorization*. 3rd ed. Oxford: Oxford University Press.

Part I

Basic Questions