Preface

Andreea and Laurie

This book is not an English grammar: it makes no attempt to cover everything you might want to know about the way English works as a language. Rather it is a book about aspects of the grammar of English: it picks out some of the interesting parts and stirs up the mud at the bottom of the linguistic pond in its search for gems. It tries to explain why people might not know what is the right thing to say or to write (beyond mere ignorance), and what keeps different versions of a particular grammatical pattern in use. It also tries to point out some of the complexities of grammar that we all deal with on a daily basis. Unlike other books of grammar (though not all!), our book provides details of the language and the patterns observed but also about the people who search for these patterns (linguists). In some chapters, we illustrate problems and issues which researchers are struggling to solve as well as the methods they use to do so, unveiling the process by which we study language and difficulties which our methods present us with. Knowledge about grammar is an emerging and ongoing field of inquiry, and by no means a completed task.

In order to do this, we have to look at the way people really use language – what they actually say and write as opposed to what they might think they say or write (which is different again from what somebody else might think they ought to say and write). Because we both live in New Zealand, we have often used brief excerpts from New Zealand media for illustrative purposes. We cite quality New Zealand newspapers like the *Dominion Post* and the *New Zealand Herald*, we cite the prestigious RNZ National (Radio New Zealand), especially in news and current affairs programmes. This is simply because the material is readily available to us. We do not believe that the language of the New Zealand media is any different in relevant ways from the language of the media in Australia or the United Kingdom (and although there are sometimes marked differences from the language of North America, we point out where this may be relevant to our discussion).

We would like to thank Winifred Bauer, Jessie Burnette, Paul James and David Trye for their readings of drafts, enthusiastic support and general help. We are also greatly indebted to Routledge editors, especially Louisa Semlyen and Eleni Steck for their belief in our ideas, and the referees for their thoughtful comments and suggestions.

1 Introduction

Laurie and Andreea

Setting the scene

English is currently the most described language in the world. We say this not only because of the number of very detailed grammars of English such as Jespersen (1909–1949), Quirk et al. (1985) or Huddleston & Pullum (2002), and not only because ~~of~~ we have been writing grammars of English for a very long time; one of the earliest works being Bullokar's *Pamphlet for Grammar* in 1586 (Linn 2006, p. 74). While both these factors are clearly important, it is the range of constructions which are part of English that have been subject to detailed consideration by generations of grammarians and linguists, either trying to provide a better description of English, or trying to prove some theoretical point about the way grammar works. The descriptions do not always agree (for various reasons, some of which we will consider here), but they are available for consultation, and they provide a very detailed description of more of the grammar of English than has been attempted for any other language. It might thus seem that yet another book about English grammar would be a waste of time. Surely, it has all been done, and the job of English grammarians has now been rendered redundant.

In this book, we argue that this is far from being the case. Researchers who work in the field of English grammar are still making new discoveries, finding new regularities and providing new insights. There are many reasons why we might not know precisely what is going on in grammar, not only in English, but in other languages, too.

Variation in language

First, we should note that English is not a homogeneous language. There are many varieties of English, where 'variety' is a technical term that encompasses dialects, styles, genres, even patterns which may be indicative of the individual speaker or writer. The examples given in (1)–(3) are ways of saying the same thing in different varieties.

(1) a. I've not finished it.
    b. I haven't finished it.

DOI: 10.4324/9781003148999-1

(2) a. I don't know whom to trust.
    b. I don't know who to trust.
(3) a. I have a friend lives in Auckland.
    b. I have a friend who lives in Auckland.

In the United Kingdom, (1a) is more likely to be heard in the north (Scotland and northern England), (1b) in the south; (2a) is much more formal than (2b), with the result that (2b) is more likely to occur in spoken language and (2a) is often confined to some kinds of written language (Microsoft Word suggests changing (2b)). The sentence in (3a) illustrates a structure which can be heard, but is not generally accepted, and which is found more often in informal styles and in some dialect areas, while (3b) is the standard form (the form of the language which is deemed most prestigious and used in formal contexts). Different varieties may use different grammatical patterns (they may also be different varieties because they sound different – in Lancashire, England, *wood* may rhyme with *cud* – or use different words altogether – *slater* and *woodlouse* are both words to denote the same creature), so that a person from Perth, Australia, may use different grammatical patterns from a person from Perth, Scotland, a lawyer may use different grammatical patterns (speak a different variety) from a carpenter, a man may use different grammatical patterns from a woman, a woman may use different grammatical patterns in addressing her daughter and in addressing her grandmother, and so on.

One particular kind of variety is brought about by language change: things which used to be normal in English are now no longer normal. In the late 1800s, (4a) was the ordinary way of expressing the meaning, and (4b) did not exist, though today speakers would probably find (4a) very odd (see Strang 1970, pp. 98–9).

(4) a. The house is building.
    b. The house is being built.

Language changes all the time, and once the time-depth is great enough, the old seems odd and the new seems normal. But while the change is actually occurring and both forms can be found, change can cause some confusion. Such periods of confusion can sometimes be very long, and then we often find that the speech of younger speakers is systematically different from that of older speakers. This confusion sometimes gives rise to statements of what 'should' be done, which can have the effect of prolonging the life of a moribund construction.

Descriptive and prescriptive grammar

Notions of what should be done in grammar are called 'prescriptive' or 'normative' statements. Prescriptive ideas, ideas about the way language ought to be, have several origins. In some cases, a grammarian will describe what he or she believes is the case in the language of people who are thought to use the 'best' English – often literary writers, biographers and historians – and this gets interpreted as meaning that everyone should copy this usage.

In some instances, the copied usage is a minority one. For example, it is probably the case that for the majority of English speakers around the world (though more often in some areas than in others), *I never did it last week* is a perfectly normal statement, even though, in standard English, *never* is incompatible with *last week*, and *I didn't do it last week* would be the expected form.

In other cases, it may be suggested that English grammar should follow a pattern that is found in some other language (a language with high cultural prestige, often Greek or Latin). For example, the class of prepositions is so-called because in Latin, the corresponding set of words occurred before (pre-) noun phrases. In English, they are also found in that position, but the same forms also have another use in expressions like *come out*, *drop in*, *put up*, and in other places where they do not always occur before a noun phrase. Lowth (1762) knew that what we now call 'preposition stranding' was 'an Idiom which our language is strongly inclined to' [sic] (quoted in Huddleston & Pullum 2020, p. 202), and it was later authors who tried to impose what they perceived as a Latin pattern on English and left us with awkward normative statements such as 'Never end a sentence with a preposition' (to which, of course, the response may be 'What would I want to end a sentence with a preposition for?' or 'What would I want to use a preposition to end a sentence up with for?').

In a third set of cases, grammarians may describe an older pattern and readers may take it as implying (whether that is the author's intention or not) that the newer pattern replacing it should not be used (as is illustrated with *whom* in (2a)).

By contrast, so-called 'descriptive' statements about language claim to do no more that state what grammatical patterns can be observed in usage. They may also try to provide explanations of the observed patterns, and the more modern grammars may indicate where there are differences of opinions on such matters. The three major grammars cited above are all descriptive in this sense, but all descriptive works may end up being used normatively, for example in second-language teaching situations.

What becomes of interest then, is how people know what general usage is, and how to identify appropriate data for description. Many writers make up their examples (as we have done to this point), and assume that their readers will agree with them about what is or is not part of English. In fundamental sentences like (5), that may well be true, but there are plenty of instances where we might disagree, and the disagreement may be minor or major.

(5) The cat sat on the mat.

For example, in the third decade of the twenty-first century there is likely to be disagreement about the sets in (6)–(7); even if individual speakers are consistent in their own usage, there may be disagreement between speakers.

(6) a. I'm bored with this game.
    b. I'm bored of this game.
(7) a. Have you any money?
    b. Do you have any money?
    c. Have you got any money?

Sources of evidence

To avoid the bias of citing one's own personal intuition – which, in some cases, may change from day to day – many grammarians prefer to cite examples of usage from users who have, as far as we can tell, not been consciously manipulating their usage when they produced the relevant material (and we do this quite often in this book). It is possible, of course, for people to make mistakes in their usage – or at least, to say things which they would wish to correct if they had more time. Making mistakes is particularly easy to do when speaking as opposed to writing, but it can happen in any language medium. To avoid such errors, we would ideally like to see a particular pattern repeated from several authors, but it is often difficult to find suitable examples. To maximize the chances of getting good data, linguists and grammarians these days frequently use corpora.

A corpus is a large body of data selected in such a way as to provide usage from many people, and entered into a computer database to allow for easy searching. Corpora can be anything from a few thousand words of running text to billions of words of running text, depending on when the texts in the corpus were collected and how they were gathered and treated. Typically, for example, corpora of spoken English are smaller than corpora of written English, because a lot more work is involved in entering and editing data for a corpus of spoken language than for one of written language. Older corpora (from the late twentieth century), which took material from printed media, tend to be smaller than modern corpora which take material from internet sources. We provide examples from a number of corpora in this book, and the corpora used are listed in our set of references, along with the books that we have consulted.

Some corpora are parsed, which means that each word is assigned to a word class, such as noun, verb, adjective and so on. Because most of the parsing is done automatically, using artificial intelligence and statistical techniques, there are to date still problems involved in doing this. Though it is rarely completely accurate, in principle it makes it easy to distinguish between *standard* ('a flag') and *standard* ('normal'), or to see what nouns fit into the noun slot in *a high* [Noun] *for* (for example, we can have *high hopes/regards/praise for*, but not, *high corpus for*). While corpora allow us access to a great deal of good data, the data has been democratized, in the sense that the data in such collections is typically the written usage of everyday writers, with no priority given to writers who are considered models of the best usage, as was typically the case in earlier descriptions, including dictionaries.

What is grammar?

We have not yet considered the meaning of the word *grammar*, a word that has an unfortunately large number of meanings among linguists and grammarians (Crystal 2015, Bauer 2021).

In one usage, grammar is the science of the study of all language phenomena. This might involve phonetics (is the [s] sound in *cuts* in *She cuts the grass every*

week systematically different from the [s] sound in *cuts* in *She has several bad cuts on her hand* ?), phonology (are the rules for stressing nouns different from those for stressing verbs in English?) and pragmatics (can you give an order by making a statement?) – incidentally, the answers to all of these questions is 'yes'.

A more restrictive use of grammar omits these factors, and keeps the word 'grammar' for two things: syntax (the way in which words are ordered and how that relates to the meaning of larger units) and morphology (the way in which meaningful elements go together to make up words, so that we can have *wasp-ish-ness* but not **wasp-ish-ity* or **wasp-ity-ish*). (The asterisk is a device conventionally used by linguists to show that what follows is not good English.)

This latter is the main sense in which we will use the word. *A grammar* (not just *grammar* any more) will also be used, as it has already been used in the second example, to denote a description of a language or a part of a language focusing on these questions.

Grammars, in this sense, focused heavily on regularities in the language. Their interest was to uncover and document regular patterns in the language. A typical example of such a regular pattern might be how verbs are adjusted for different persons (first, second, third). The left-hand side in (8) gives the Latin version of the how the verb meaning 'love' is 'conjugated' and the right-hand side gives the English version.

(8) a. amo 'I love'
b. amas 'you love' (just one person)
c. amat 'he/she/it loves'
d. amāmus 'we love'
e. amātis 'you love' (more than one person)
f. amant 'they love'

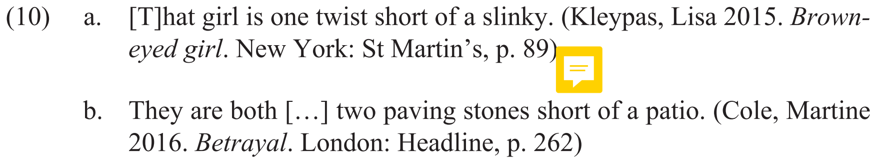
You can immediately spot that Latin makes a lot of changes along the way, while English does not. Linguists term these 'paradigms', for instance the above pattern shows the present tense verb paradigm in Latin and English. We can show other paradigms too, such as, different tenses in English, as in (9).

(9) a. infer (basic form, infinitive)
b. infers (third person singular present tense form)
c. inferred (past tense form, past participle form)
d. inferring (present participle form)

On the basis of this paradigm, we can then predict that the past tense and past participle form corresponding to *travel* will be *travelled* (as long as *travel* belongs to the same paradigm as *infer*; *think* does not).

Paradigms can involve regularities of single words, like (8) or (9), or they can be extended to groups of words (phrases) or larger combinations of words. This kind of approach to grammar is now called a 'constructional' approach.

Constructions are patterns of words which recur with a fixed meaning. Sometimes constructions are very specific (as with the examples in (10)), and sometimes they are very general (as with the examples in (11)). In either case, a certain amount of substitution is permitted within the general meaning assigned to the construction.



(10) a. [T]hat girl is one twist short of a slinky. (Kleypas, Lisa 2015. *Brown-eyed girl*. New York: St Martin's, p. 89)

b. They are both […] two paving stones short of a patio. (Cole, Martine 2016. *Betrayal*. London: Headline, p. 262)

(11)					
	a.	My uncle	gave	my sister	a doll.
	b.	Father Christmas	gave	me	a jigsaw puzzle.
	c.	The cat in the hat	gave	us all	a big shock.
	d.	I	showed	my wife	the report.

In (10) we have a construction which, whatever words are used, means 'not very clever'; in (11) we have what is sometimes called 'the double object construction', where certain verbs can take two objects (direct and indirect, the thing given and the person it is given to).

A very different metaphor is the notion of a descriptive rule. A descriptive rule is a rule which tells you what is happening. Depending on the work you happen to be reading, it may be formulated in anything from ordinary prose (e.g. 'The indirect object precedes the direct object' in (11)) to a formula whose elements are theoretically determined (e.g. 'V NP NP', where 'V' is short for 'verb' and 'NP' is short for 'noun phrase', for the same sentences in (11)). More modern formal grammars tend to take the latter approach.

How many answers are there to a question?

A question such as *Which side of the road do they drive on in Japan?* has a single and verifiable answer. People drive on the left in Japan, and you can go to Japan and see for yourself that this is the case. A question like *Why did people vote for Brexit in the referendum?*, on the other hand, has multiple answers. We can easily list a few: they were misled by propaganda; they were not deflected from important issues by the arguments presented in the anti-Brexit propaganda; they wanted to control immigration; they wanted Britain to be great again; they could not imagine ever trying to get a job in Europe; they did not think that it was important that a country should keep its word on alliances; they did not understand the economic advantages of the EU; they thought the EU was overriding their own democracy. You can probably add to this list. The thing about a question of this type is that there is not just one right answer, and that more than one answer can be true at the same time. Also, the answer may be different for different subgroups of people.

Questions about language are more often like the Brexit question than like the side-of-the-road question. There may be multiple answers, people may disagree

about which of them are important, different people may behave as they do due to different sets of answers. If you ask a question such as

(12) Is it *She went to lay on the bed* or *She went to lie on the bed*?

there is not necessarily a single right answer. Who you are and who you are speaking to, what message you want to convey, where you come from and whether or not you are attempting to speak standard English may all be relevant factors in an answer. On the other hand, a question such as

(13) Is it *a stationery car* or *a stationary car*?

may have a single answer, provided that you make assumptions about expecting standard spellings and standard situations.

Multiple answers, or the inability to give a single, simple answer, should not be a sign that the question is silly or that the person expected to answer does not know what they are talking about. The lack of a unique answer is implicit in some questions, and where the answer depends on identity, personality or a myriad of social factors, the lack of a single answer is not a cop-out. The chapters in this book are not intended to provide definitive answers to the questions they raise. Nor are they intended to provide an exhaustive list of every question raised in connection to the topic of the chapter at hand. Our goal is to take readers on a grammatical journey into the kinds of problems that those who study grammar (and particularly for us, English grammar) ask. We would like to provide readers with a chance to think about language – and possibly to think about language in a new way.

Why are there grammar mysteries?

Finally, we come to the notion of a mystery. There are several reasons why a particular piece of usage might be considered a mystery:

- We don't know what is going on; perhaps we are unable to determine any regular pattern – things are messy; perhaps we can see some regularities, but do not understand what drives the patterns we find; perhaps we just do not yet know what the relevant patterns are. If there are genuinely no patterns, the system is presumably unknowable, so this is a situation we do not expect to find, and do not want to find. In such cases, we must try to find something that makes it in principle possible to learn the grammar. There is a big difference between the unknown and the unknowable.
- We know what is going on, but it does not seem to be predictable. This might be because speakers can manipulate the patterns in subtle ways that we cannot fully discern.
- We know what is going on, but we do not know how the mind determines what will actually be said.

- We know more or less what is going on, but we do not know how to successfully capture the patterns we see within a neat theoretical framework (a grammatical model).

The last of these questions has exercised the minds of a lot of linguists over the last half-century, and although the discussions are frequently fascinating, they tend to be rather technical. Accordingly, we shall not deal in detail with the last point here, and to the extent that we do, it will just be in general terms. This should not be taken to mean that we do not care about these models. If we are really going to teach computers to use human language, it will involve creating such models that can be interpreted by artificial intelligence. It is just that, for our purposes, choosing a particular model and introducing the technicalities it involves would interfere with our purpose of drawing attention to the puzzles which are at the core of what we are interested in.

Put simply, this book is about some of the bits of English grammar that we do not understand properly. This is odd. At one level, we must surely understand the grammar well enough to use it on a daily basis: we find it easy to say what we want to say, but very difficult to introspect about how we go about putting these utterances together. And there must be regularities ('rules' if you wish) involved. If there were not, we could never make any mistakes (because a mistake implies that we have done something wrong, and if there are no rules or regularities, there is no right or wrong). Yet we all know that we are perfectly capable of making mistakes when we talk and write, and text, even in our native language. If there were no rules, moreover, foreigners would not be able to get things wrong when they talk in our language, but we do recognize typical 'foreigner' errors (*I knowed the answer, He go home early*). We also know that poets and advertising people are capable of extending our language and saying things in ways that have never been said before (*Eat Fresh, Because you're worth it*). What is the difference between extending our usage and making a mistake? Is it a matter of degree, or is there a difference of type?

Who is this book for and how should it be read?

We wrote this book because we find English grammar fascinating and because we are convinced that much of what people like us find interesting will be equally interesting for those who are not familiar with the field of linguistics. Much of language study is hidden behind fancy terms and paywall academic journals but we feel that at least some of this content should be made accessible for a wide audience.

Hence our book is intended for a non-specialist public, for English users and speakers, for advanced learners of English as a second or foreign language, for language enthusiasts and grammar fanatics – for anyone with an interest in and curiosity about language.

The chapters are not connected to each other, although their arrangement was decided upon on the basis of complexity to some extent. Regardless, readers can

pick and choose the order in which they read each chapter. The list of chapters is not exhaustive – there are many more English grammatical bits that we could have discussed, but we chose a list of those constructions that we, as grammarians, find interesting and perhaps straightforward to write about for a non-specialist audience. We have deliberately tried to make the language jargon-free wherever possible, and only introduced terms if we deemed it absolutely necessary. Each term is explained in the relevant chapter, but for convenience, we also include a chapter at the end of the book with some short definitions and examples for easy reference (see Glossary).

Wrapping up

The questions and mysteries we raise in the chapters of the book are in many ways neither completely new (some have been the focus of many research articles), nor entirely problematic (despite the gaps in current understanding of English grammar, we can still all get by using the language to communicate with one another). As language researchers, we find these questions not just captivating but also of practical importance. In the end, if we do not know how English grammar works in detail, how are we supposed to teach foreigners our language, and how can we edit texts so that they use a standard form of the language? Some people may not care: as long as we can understand one another, all is well. Others care very deeply. But whether you care or not, there is a fascinating intellectual puzzle here: what is going on, and how must the mind work to deal with the complexities of our language? If, as some have argued (Simpson 1994, p. 1894; Chomsky 1972, p. 100; Fry 2019) language makes us human, trying to see how language works is probing the very essence of our humanity.

References

Bauer, Laurie 2021. *The linguistics student's handbook*. 2nd edition. Edinburgh: Edinburgh University Press.

Bullokar, William 1586. *Pamphlet for grammar*. London: Denham.

Chomsky, Noam 1972. *Language and mind*. Enlarged edition. New York: Harcourt Brace Jovanovich. (For the 3rd edition: DOI:10.1017/CBO9780511791222)

Crystal, David 2015. *A dictionary of linguistics and phonetics*. 6th edition. Malden, MA: Blackwell. doi:10.1002/9781444302776

Fry, Stephen 2019. Interview on the BBC, 3 July.

Huddleston, Rodney & Geoffrey K. Pullum (eds) 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. doi:10.1017/9781316423530

Huddleston, Rodney & Geoffrey K. Pullum 2020. Modern and traditional descriptive approaches. In Bas Aarts, Jill Bowie & Gergana Popova (eds), *The Oxford handbook of English grammar*, 201–221. Oxford: Oxford University Press. doi:10.1093/oxfordhb/9780198755104.013.15

Jespersen, Otto 1909–1949. *A modern English grammar on historical principles*. Seven vols. London: Allen and Unwin. doi:10.4324/9780203715987

Linn, Andrew 2006. English grammar writing. In Bas Aarts & April McMahon (eds), *The handbook of English linguistics*, 72–92. Malden, MA: Blackwell.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik 1985. *A comprehensive grammar of the English language*. London: Longman.
Simpson, J.M.Y. 1994. Language. In R.E. Asher (ed.), *The encyclopedia of language and linguistics*. Vol 4. Oxford: Pergamon.
Strang, Barbara M.H. 1970. *A history of English*. London: Methuen.