

# Preliminaries

GEOFFREY K. PULLUM  
RODNEY HUDDLESTON

- 1 The aim of this book** 2
- 2 Prescriptivism, tradition, and the justification of grammars** 5
  - 2.1 Prescriptive and descriptive approaches: goals and coverage 5
  - 2.2 Disagreement between descriptivist and prescriptivist work 6
- 3 Speech and writing** 11
  - 3.1 The representation of English pronunciation 13
    - 3.1.1 Rhotic and non-rhotic accents 13
    - 3.1.2 An accent-neutral phonological representation 14
  - 3.2 Pronunciation and spelling 17
- 4 Theoretical framework** 18
  - 4.1 Description and theory 18
  - 4.2 Basic concepts in syntax 20
    - 4.2.1 Constituent structure 20
    - 4.2.2 Syntactic categories 21
    - 4.2.3 Grammatical constructions and functions 23
  - 4.3 Morphology, inflectional and lexical 26
  - 4.4 Defining grammatical concepts 28
- 5 Semantics, pragmatics, and meaning relations** 33
  - 5.1 Truth conditions and entailment 34
  - 5.2 Non-truth-conditional aspects of sentence meaning 36
  - 5.3 Pragmatics and conversational implicatures 37
  - 5.4 Pragmatic presupposition 40

## 1 The aim of this book

This book is a description of the grammar of modern Standard English, providing a detailed account of the principles governing the construction of English words, phrases, clauses, and sentences. To be more specific, we give a synchronic, descriptive grammar of general-purpose, present-day, international Standard English.

### ■ Synchronic versus diachronic description

A **synchronic** description of a language is a snapshot of it at one point in time, the opposite of a **diachronic** or historical account. English has a rich history going back over a millennium, but it is not the aim of this book to detail it. We include only a few notes on historical points of interest that will assist the reader to understand the present state of the language.

Of course, at any given moment English speakers with birthdates spread over about a century are alive, so the idea of English as it is on one particular day is a fiction: the English used today was learned by some speakers at the end of the twentieth century and by others near the beginning. But our practice will be to illustrate relevant points mainly with examples of use of the language taken from prose produced since the mid twentieth century. Examples from earlier periods are used only when particularly apposite quotations are available for a point on which the language has not subsequently changed. Wherever grammatical change has clearly occurred, our aim will be not to describe the evolutionary process but rather to describe the current state of the language.

### ■ Description versus prescription

Our aim is to **describe** and not **prescribe**: we outline and illustrate the principles that govern the construction of words and sentences in the present-day language without recommending or condemning particular usage choices. Although this book may be (and we certainly hope it will be) of use in helping the user decide how to phrase things, it is not designed as a style guide or a usage manual. We report that sentences of some types are now widely found and used, but we will not advise you to use them. We state that sentences of some types are seldom encountered, or that usage manuals or language columnists or language teachers recommend against them, or that some form of words is normally found only in informal style or, conversely, is limited to rather formal style, but we will not tell you that you should avoid them or otherwise make recommendations about how you should speak or write. Rather, this book offers a description of the context common to all such decisions: the linguistic system itself.

### ■ General-purpose versus special-purpose

We exclude from consideration what we refer to as **special-purpose** varieties of the language. Newspaper headlines, road signs, notices, and the like have their own special styles of abbreviation (*Man bites dog, arrested; EXIT ONLY THIS LANE*), and we do not provide a full treatment of the possibilities. Likewise, we do not provide a description of any special notations (chemical formulae, telephone numbers, email addresses) or of the special language found in poetry, heraldic descriptions, scientific works, chemical compound naming, computer jargon, mathematical proofs, etc. To some small extent there may be idiosyncratic grammatical patterns found in such areas, but we set them aside, avoiding complicated digressions about usages found within only a very narrow range of discourse.

### ■ Present-day English versus earlier stages

Modern English is generally defined by historians of English to be the English used from 1776 onwards. The recent part of the latter period (say, since the Second World War) can be called Present-day English. Linguistic changes have occurred in the grammar of English during the Modern English period, and even during the last half-century. Our central aim is to describe Present-day English in its standard form. This means, for example, that we treat the pronoun system as not containing a contrast between familiar and respectful 2nd person pronouns: the contrast between *thou* and *you* has been lost, and we do not mention *thou* in this grammar. Of course, this does not mean that people who use *thou* (actors in period plays, people addressing God in prayers, or Quakers who have retained the older usage) are making a mistake; but they are not using the general-purpose standard Present-day English described in this book.

### ■ Grammar versus other components

A **grammar** of a language describes the principles or rules governing the form and meaning of words, phrases, clauses, and sentences. As such, it interacts with other components of a complete description: the **phonology** (covering the sound system), the **graphology** (the writing system: spelling and punctuation), the dictionary or **lexicon**, and the **semantics**.

Phonology and graphology do not receive attention in their own right here, but both have to be treated explicitly in the course of our description of inflection in Ch. 18 (we introduce the concepts that we will draw on in §3 of this chapter), and Ch. 20 deals with one aspect of the writing system in providing an outline account of the important system of punctuation.

A lexicon for a language deals with the vocabulary: it brings together information about the pronunciation, spelling, meaning, and grammatical properties of the **lexical items** – the words, and the items with special meanings that consist of more than one word, the idioms.

The study of conventional linguistic meaning is known as **semantics**. We take this to cut across the division between grammar and lexicon. That is, we distinguish between **lexical semantics**, which dictionaries cover, and **grammatical semantics**. Our account of grammatical meaning will be quite informal, but will distinguish between semantics (dealing with the meaning of sentences or words as determined by the language system itself) and **pragmatics** (which has to do with the use and interpretation of sentences

as used in particular contexts); an introduction to these and other concepts used in describing meaning is given in §5 of this chapter.

A grammar itself is divisible into two components, **syntax** and **morphology**. Syntax is concerned with the way words combine to form phrases, clauses, and sentences, while morphology deals with the formation of words. This division gives special prominence to the **word**, a unit which is also of major importance in the lexicon, the phonology and the graphology.

### ■ Standard versus non-standard

Perhaps the most subtle concept we have to rely on is the one that picks out the particular variety of Present-day English we describe, which we call Standard English. Briefly (for we will return to the topic below), we are describing the kind of English that is widely accepted in the countries of the world where English is the language of government, education, broadcasting, news publishing, entertainment, and other public discourse.

In a large number of countries (now running into scores), including some where most of the people have other languages as their first language, English is used for most printed books, magazines, newspapers, and public notices; for most radio and television broadcasting; for many or most film scripts, plays, poetry, and other literary art; for speeches, lectures, political addresses, proclamations, official ceremonies, advertisements, and other general announcements. In these countries there is a high degree of consensus about the appropriate variety of English to use. The consensus is confirmed by the decisions of broadcasting authorities about the kind of English that will be used for public information announcements, newscasts, commentaries to broadcasts of national events such as state funerals, and so on. It is confirmed by the writing found in magazines, newspapers, novels, and non-fiction books; by the editing and correcting that is done by the publishers of these; and by the way writers for the most part accept such editing and correcting of their work.

This is not to say that controversy cannot arise about points of grammar or usage. There is much dispute, and that is precisely the subject matter for prescriptive usage manuals. Nonetheless, the controversy about particular points stands out against a backdrop of remarkably widespread agreement about how sentences should be constructed for such purposes as publication, political communication, or government broadcasting. This widespread agreement defines what we are calling Standard English.

### ■ National versus international

Finally, we note that this book is not intended to promote any particular country's variety of Standard English as a norm; it is to apply internationally. English is the single most important language in the world, being the official or *de facto* language of the United Kingdom, the United States of America, Canada, Australia, New Zealand, South Africa, and dozens of others, and being the *lingua franca* of the Internet. Many varieties of English are spoken around the world – from lectures in graduate schools in Holland to parliamentary proceedings in Papua New Guinea – but interestingly the vast majority of the variation lies in pronunciation and vocabulary. The number of differences in grammar between different varieties of Standard English is very

small indeed relative to the full range of syntactic constructions and morphological word-forms.

Nevertheless, there undoubtedly are differences of this kind that need to be noted. For example, the use of the verb *do* following an auxiliary verb, as in <sup>%</sup>*I'm not sure that I'll go, but I may do* is not found in American English, and conversely the past participle verb-form *gotten*, as in <sup>%</sup>*I've just gotten a new car*, is distinctively American. We use the symbol ‘%’ to mark constructions or forms that are restricted to some dialect or dialects in this way.

The regional dialects of Standard English in the world today can be divided into two large families with regional and historical affinities. One contains standard educated Southern British English, henceforth abbreviated **BrE**, together with a variety of related dialects, including most of the varieties of English in Great Britain, Australia, New Zealand, South Africa, and most other places in the British Commonwealth. The second dialect family we will refer to as American English, henceforth **AmE** – it contains the dialects of the United States, Canada, and associated territories, from Hawaii and Alaska to eastern Canada.

## 2 Prescriptivism, tradition, and the justification of grammars

The topic of prescriptivism and its relation to the long tradition of English grammatical scholarship needs some further discussion if the basis of our work, and its relation to other contributions to the field, is to be properly understood. It relates to the issue of how the statements of a grammar are justified: what the support for a claimed grammatical statement might be.

### 2.1 Prescriptive and descriptive approaches: goals and coverage

The distinction between the prescriptive and descriptive approaches to grammar is often explained by saying that prescriptivists want to tell you how you **ought** to speak and write, while descriptivists want to tell you how people actually **do** speak and write. This does bring out the major difference between the two approaches: it is a difference in goals. However, it is something of an oversimplification, because writing a descriptive grammar in practice involves a fair amount of idealisation: we need to abstract away from the errors that people make, especially in speech (this point is taken up again in §3 below). In addition, it glosses over some significant differences between the kinds of works prescriptivists and descriptivists characteristically produce.

#### ■ Differences in content

The basic difference in goals between prescriptive and descriptive works goes hand in hand with a striking difference in topics treated. The subject matters overlap, but many topics dealt with by prescriptive works find no place in a descriptive grammar, and some topics that must be treated in a descriptive grammar are universally ignored by prescriptive works.

The advice of prescriptivists is supplied in works of a type we will refer to as **usage manuals**. They are almost invariably arranged in the style of a dictionary, containing an

alphabetically arranged series of entries on topics where the issue of what is correct or acceptable is not altogether straightforward. In the first few pages of one usage manual we find entries on *abacus* (should the plural be *abaci*?), abbreviations (which ones are acceptable in formal writing?), *abdomen* (is the stress on the second syllable or the first?), *abduction* (how does it differ in meaning from *kidnapping*?), and so on. These points concern inflection, formal writing, pronunciation, and meaning, respectively, and on all of them a degree of variation and occasional uncertainty is encountered even among expert users of English. Not all of them would belong in a grammatical description. For example, our grammar does cover the plural of *abacus* (Ch. 18, §4.1.6), but it does not list abbreviations, or phonological topics like the placement of stress in English words, or lexical semantic topics like the distinction between *abduction* and *kidnapping*. These we take to be in the province of lexicon – matters for a dictionary rather than a grammar.

Usage manuals also give a great deal of attention to matters of style and effective expression that lie beyond the range of grammar as we understand it. Thus one prescriptive usage dictionary warns that *explore every avenue* is a tired cliché (and adds that it makes little sense, since exploration suggests a more challenging environment than an avenue); that the phrase *in this day and age* ‘should be avoided at all costs’; that *circling round* is tautologous (one can only circle by going round) and thus should not be used; and so on. Whether or not one thinks these are good pieces of advice, we do not take them to fall within the realm of grammar. A sentence like *In this day and age one must circle round and explore every avenue* may be loaded with careworn verbiage, or it may even be arrant nonsense, but there is absolutely nothing **grammatically** wrong with it.

There are also topics in a descriptive grammar that are uniformly ignored by prescriptivists. These include the most salient and well-known principles of syntax. Prescriptive works tend to be highly selective, dealing only with points on which people make mistakes (or what are commonly thought to be mistakes). They would never supply, for example, the grammatically important information that determinatives like *the* and *a* precede the noun they are associated with (*the house*, not *\*house the*),<sup>1</sup> or that modal auxiliaries like *can* and *must* are disallowed in infinitival clauses (*\*I'd like to can swim* is ungrammatical), or that in subordinate interrogative clauses the interrogative element comes at the front (so we get *She asked what we needed*, not *\*She asked we needed what*). Native speakers never get these things wrong, so no advice is needed.

## 2.2 Disagreement between descriptivist and prescriptivist work

Although descriptive grammars and prescriptive usage manuals differ in the range of topics they treat, there is no reason in principle why they should not agree on what they say about the topics they both treat. The fact they do not is interesting. There are several reasons for the lack of agreement. We deal with three of them here: (a) the basis in personal taste of some prescriptivist writers' judgements; (b) the confusion of informality with ungrammaticality; and (c) certain invalid arguments sometimes appealed to by prescriptivists. These are extraneous features of prescriptive writing about language rather than inherent ones, and all three of them are less prevalent now than they were

<sup>1</sup>Throughout this book we use an asterisk to indicate that what follows is ungrammatical.

in the past. But older prescriptive works have exemplified them, and a few still do; their influence lingers on in the English-speaking educational world.

### (a) Taste tyranny

Some prescriptivist works present rules that have no basis in the way the language is actually used by the majority of its native speakers, and are not even claimed to have any such basis – as though the manual-writer’s own judgements of taste took precedence over those of any other speaker of the language. They expect all speakers to agree with their judgements, no matter what the facts of language use might show.

For example, one usage manual, discussing why it is (supposedly) incorrect to say *You need a driving instructor who you have confidence in*, states that ‘The accusative *whom* is necessary with the preposition *in*, though *whom* is a word strangely shunned by most English people.’ We take the implication to be that English people should not shun this word, since the writer (who is English) does not. But we are inclined to ask what grounds there could be for saying that *whom* is ‘necessary’ if most English people (or speakers of the English language) would avoid it.

The same book objects to *centre (a)round*, calling it incorrect, although ‘probably more frequently used than the correct *centre on*’. Again, we wonder how *centre (a)round* can be determined to be incorrect in English if it is indeed more commonly used by English speakers than what is allegedly correct. The boundary would appear to have been drawn in the wrong place.

Prescriptive works instantiating this kind of aesthetic authoritarianism provide no answer to such obvious questions. They simply assert that grammar dictates things, without supporting their claim from evidence. The basis for the recommendations offered appears to lie in the writer’s taste: the writer quoted above simply does not like to see *who* used where it is understood as the object of a preposition, and personally hates the expression *centre around*. What is going on here is a universalising of one person’s taste, a demand that everyone should agree with it and conform to it.

The descriptivist view would be that when most speakers use a form that our grammar says is incorrect, there is at least a *prima facie* case that it is the grammar that is wrong, not the speakers. And indeed, even in the work just quoted we find the remark that ‘*Alright* is common, and may in time become normal’, an acknowledgement that the language may change over time, and what begins as an isolated variant on a pattern may eventually become the new pattern. The descriptive grammarian will always adopt a stance of something more like this sort, thus making evidence relevant to the matter at hand. If what is involved were a matter of taste, all evidence would be beside the point. But under the descriptive viewpoint, grammar is not a matter of taste, nor of aesthetics.

This is not to say that the expression of personal aesthetic judgements is without utility. The writer of a book on usage might be someone famous for brilliant use of the language, someone eminently worthy of being followed in matters of taste and literary style. It might be very useful to have a compendium of such a person’s preferences and recommendations, and very sensible for a less expert writer to follow the recommendations of an acknowledged master of the writer’s craft (assuming such recommendations do reliably accord with the master’s practice). All we are pointing out is that where the author of an authoritarian usage manual departs from recommendations that agree with the way most people use the language, prescriptivist and descriptivist



belongs to (very) formal style, while accusative *me* is neutral or informal (again, see Ch. 5, §16.2.1 for a fuller description of the facts). Confusing informality with ungrammaticality again, a strong prescriptivist tradition says that only [2a] is grammatical. The accusative *me* is claimed to be the case of the direct object, as in *It hurt me*, but in [2] the noun phrase after the verb is a predicative complement. In Latin, predicative complements take nominative, the same case as the subject. An assumption is being made that English grammar too requires nominative case for predicative complements. Use of the accusative *me* is regarded as a departure from the rules of grammar.

The mistake here, of course, is to assume that what holds in Latin grammar has to hold for English. English grammar differs on innumerable points from Latin grammar; there is no reason in principle why the assignment of case to predicative complements should not be one of them. After all, English is very different from Latin with respect to case: the nominative–accusative contrast applies to only a handful of pronouns (rather than to the full class of nouns, as in Latin). The right way to describe the present situation in Standard English (unlike Latin) is that with the pronouns that have a nominative–accusative case distinction, the choice between the cases for a predicative complement noun phrase varies according to the style level: the nominative is noticeably formal, the accusative is more or less neutral and always used in informal contexts.

Another kind of illegitimate argument is based on analogy between one area of grammar and another. Consider yet another construction where there is variation between nominative and accusative forms of pronouns:

- [3] a. *They invited me to lunch.*                      b. <sup>%</sup>*They invited my partner and I to lunch.*

The ‘%’ symbol is again used to mark the [b] example as typically used by some speakers of Standard English but not others, though this time it is not a matter of regional variation. The status of the construction in [b] differs from that of *It’s me*, which is undisputedly normal in informal use, and from that of <sup>1</sup>*Me and Kim saw her leave*, which is unquestionably non-standard. What is different is that examples like [b] are regularly used by a significant proportion of speakers of Standard English, and not generally thought by ordinary speakers to be non-standard; they pass unnoticed in broadcast speech all the time.

Prescriptivists, however, condemn the use illustrated by [3b], insisting that the ‘correct’ form is *They invited my partner and me to lunch*. And here again they seek to justify their claim that [3b] is ungrammatical by an implicit analogy, this time with other situations found in English, such as the example seen in [a]. In [a] the pronoun functions by itself as direct object of the verb and invariably appears in accusative case. What is different in [b] is that the direct object of the verb has the form of a coordination, not a single pronoun. Prescriptivists commonly take it for granted that this difference is irrelevant to case assignment. They argue that because we have an accusative in [a] we should also have an accusative in [b], so the nominative *I* is ungrammatical.

But why should we simply assume that the grammatical rules for case assignment cannot differentiate between a coordinated and a non-coordinated pronoun? As it happens, there is another place in English grammar where the rules are sensitive to this distinction – for virtually all speakers, not just some of them:

- [4] a. *I don’t know if you’re eligible.*                      b. <sup>\*</sup>*I don’t know if she and you’re eligible.*

The sequence *you are* can be reduced to *you're* in [a], where *you* is subject, but not in [b], where the subject has the form of a coordination of pronouns. This shows us not only that a rule of English could apply differently to pronouns and coordinated pronouns, but that one rule actually does. If that is so, then a rule could likewise distinguish between [3a] and [3b]. The argument from analogy is illegitimate. Whether [3b] is treated as correct Standard English or not (a matter that we take up in Ch. 5, §16.2.2), it cannot be successfully argued to be incorrect simply by virtue of the analogy with [3a].

The claim that [11b] (*It's clear who they had in mind*) is ungrammatical is supported by the same kind of analogical reasoning. In *They had me in mind*, we have accusative *me*, so it is assumed that the grammar likewise requires accusative *whom*. The assumption here is that the rules of case-assignment are not sensitive to the difference in the position of the pronoun (after the verb for *me*, at the beginning of the clause for *who*), or to the difference between interrogative and personal pronouns. There is, however, no basis for assuming that the rules of grammar cannot make reference to such differences: the grammar of English could assign case to clause-initial and non-clause-initial pronouns, or to interrogative and non-interrogative pronouns, in slightly different ways.<sup>2</sup>

We should stress that not all prescriptive grammarians exhibit the shortcomings we have just catalogued – universalising taste judgements, confusing informality with ungrammaticality, citing spurious external justifications, and arguing from spurious analogies. There are usage manuals that are accurate in their understanding of the facts, clear-sighted in their attitudes towards usage trends, and useful in their recommendations; such books can be an enormous help to a writer. But the good prescriptive manuals respect a crucial tenet: that their criterion should always be the use of the standard language by its native speakers.

As we have said, to some extent good usage manuals go far beyond grammar into style, rhetoric, and communication, giving advice about which expressions are over-used clichés, or fail to make their intended point, or are unintentionally ambiguous, or perpetuate an unfortunate malapropism, or any of a large number of other matters that lie beyond the scope of this book. But when it comes to points of grammar, the only legitimate basis for an absolute judgement of incorrectness in a usage manual is that what is being rejected is **not in the standard language**.

The aspects of some prescriptivist works that we have discussed illustrate ways in which those works let their users down. Where being ungrammatical is confused with merely being informal, there is a danger that the student of English will not be taught how to speak in a normal informal way, but will sound stilted and unnatural, like an inexperienced reader reading something out from a book. And where analogies are used uncritically to predict grammatical properties, or Latin principles are taken to guarantee correct use of English, the user is simply being misled.

<sup>2</sup>A further type of invalid argument that falls under the present heading confuses grammar with logic. This is illustrated in the remarkably widespread but completely fallacious claim that non-standard *I didn't see nobody* is intrinsically inferior to standard *I didn't see anybody* because the two negatives cancel each other out. We discuss this issue in Ch. 9, §6.2.

The stipulations of incorrectness that will be genuinely useful to the student are those about what is actually not found in the standard language, particularly with respect to features widely recognised as characteristic of some definitely non-standard dialect. And in that case evidence from use of Standard English by the people who speak it and write it every day will show that it is not regularly used, which means prescriptive and descriptive accounts will not be in conflict, for evidence from use of the language is exactly what is relied upon by descriptive grammars such as we present here.

The evidence we use comes from several sources: our own intuitions as native speakers of the language; the reactions of other native speakers we consult when we are in doubt; data from computer corpora (machine-readable bodies of naturally occurring text),<sup>3</sup> and data presented in dictionaries and other scholarly work on grammar. We alternate between the different sources and cross-check them against each other, since intuitions can be misleading and texts can contain errors. Issues of interpretation often arise. But always, under the descriptive approach, claims about grammar will depend upon evidence.

### 3 Speech and writing

There are significant and interesting differences between spoken and written language, but we do not regard written English as a different language from spoken English. In general, we aim to describe both the written standard variety that is encountered in contemporary newspapers, magazines, and books and the spoken standard variety that is heard on radio and television programmes in English-speaking countries.

#### ■ ‘Speaker’ and ‘utterance’ as medium-neutral terms

Most of what we say will apply equally to the spoken and written varieties of the language. As there is no non-technical term covering both one who utters a sentence in speech and one who writes a sentence, we will follow the widespread practice in linguistics of extending the ordinary sense of ‘speaker’ so as to subsume ‘writer’ – a practice that reflects the fact that speech is in important respects more basic than writing.<sup>4</sup> We likewise take ‘utterance’ to be neutral between the mediums, so that we will refer to both spoken and written utterances.

#### ■ Practical bias towards written English

Despite our neutrality between speech and writing in principle, there are at least three reasons why the reader may perceive something of a bias in this work towards data from

<sup>3</sup>The computer corpora that we have made use of are the Brown corpus of a million words of American English; the London/Oslo/Bergen (LOB) corpus of British English; the Australian Corpus of English (ACE); and the *Wall Street Journal* corpus distributed by the Association for Computational Linguistics. The British National Corpus (BNC) was only released to scholars working outside the UK after the book was in final draft. We have also drawn on a variety of other sources, including collections of our own from sources such as magazines, newspapers, plays, books, and film scripts.

<sup>4</sup>Since our discussion of sentences will very often make reference to the way they are used we will have very frequent occasion to talk of speakers, and in order to avoid repeatedly using the term ‘speaker’ we will often simply use the 1st person pronoun *I*. Given that the book has joint authorship this pronoun could not be used in reference to any specific person, and hence is available as a convenient variant of ‘the speaker’.

written English. To the extent that it is present, it stems from practical considerations rather than matters of principle. We will discuss here the three factors motivating the choices we have made.

#### Citation of forms and examples

First, we normally follow the usual practice in grammars of citing words or sentences in their written form. This is mainly a matter of practical convenience: it is much more straightforward typographically, and more widely accessible to readers, to supply examples in this form. In certain cases – as, for example, in describing the inflectional forms of verbs and nouns in Ch. 18 – it is necessary to indicate the pronunciation, and for this purpose we use the system of transcription described in §3.1.2 below. Representations in written form are given in italics, while phonological representations are enclosed in obliques.

#### Accessibility of print sources

Second, we make frequent use of genuinely attested examples (often shortened or otherwise modified in ways not relevant to the point at issue), and it is significantly easier to obtain access to suitable large collections, or corpora, of written data in a conveniently archived and readily searchable form than it is for speech.

#### Error rates in speech

Third, and most importantly, it must be acknowledged that the error content of spoken material is higher than that of written material. Those who have listened to tape recordings of spontaneous conversation are likely to have been struck by the high incidence of hesitation noises, false starts, self-corrections, repetitions, and other dysfluencies found in the speech of many people. It is not hard to see why speech contains a higher number of errors than writing. The rapid production of speech (quite often several words per second) leaves little time for reflection on construction choices or planning of sentence structure, so that at normal conversational pace slip-ups of the kind mentioned are very common. As a result, what speakers actually come out with reflects only imperfectly the system that defines the spoken version of the language. Hardly noticed by the listener, and often compensated for by virtually unconscious repair strategies on the part of the speaker, these sporadic interruptions and imperfections in speech production are inherently outside the purview of the grammarian (the discipline of psycholinguistics studies them in order to learn about the planning, production, and perception of speech). They therefore have to be screened out through judicious decision-making by a skilled native speaker of the language before grammatical description is attempted. The original speaker is not always available for the tedious editing task, and so someone else has to interpret the transcript and remove the apparent errors, which means that misunderstandings can result (word sequences that were actually due to slips might be wrongly taken to represent grammatical facts).

Written English has the advantage that its slow rate of composition has generally allowed time and opportunity for nearly all these slips and failures of execution to be screened out by the actual author of the sentence. This provides a practical reason for us to show a preference for it when selecting illustrative examples: we have very good reason to believe that what ultimately gets printed corresponds fairly closely to what the writer intended to say.

The nature of the written medium and the slower sentence-planning environment permits the construction of longer sentences than typically occur in speech, but we take this to be a matter of degree, not a matter of written English instantiating new possibilities that are completely absent from the spoken language. The basic point of most written material is that people who are ordinary native speakers of the language should read it and understand it, so the pressure will always be in the direction of keeping it fairly close to the language in which (ignoring the speech errors referred to above) ordinary people talk to each other.

Thus while we acknowledge a tendency for the exemplification in this grammar to be biased towards written English, we assume that the goal of providing a description that is neutral between spoken and written English is not an unreasonable one. Sharp divergences between the syntax of speech and the syntax of writing, as opposed to differences that exist between styles within either the spoken or the written language, are rare to the point of non-existence.

### 3.1 The representation of English pronunciation

This section provides an introduction to the system of representation we use in this book in those cases where it is necessary to indicate the pronunciation of words or word sequences. Developing a system that will be readily usable by non-specialists is by no means a trivial enterprise; English has a remarkably complex vowel system compared to most other languages, and one of the most complex patterns of fit between sound and spelling found in any language. Taken together, these facts raise some significant and unavoidable difficulties even if only one variety of English is under consideration. But an additional problem is that English is a global language with something like 400 million native speakers pronouncing the language in many different ways: pronunciation differs across the world more than any other aspect of the language.

#### 3.1.1 Rhotic and non-rhotic accents

We will use the term **accent** for varieties of a language distinguished by pronunciation, opposing it to **dialect**, which applies to varieties distinguished by grammar or vocabulary. The most important accent distinction in English concerns the sound we represent as /r/. Most speakers in the BrE family of dialects have a **non-rhotic accent**: here /r/ occurs in **pre-vocalic** position, i.e. when immediately preceding a vowel, as in *run* or *area*, but not in **post-vocalic** position, after the vowel of a syllable. For example, in a non-rhotic accent there is no /r/ in any of the words in [1] (as pronounced in isolation):

- [1] i a. *mar, bear, floor, stir, actor*      b. *care, hire, bore, sure, cure*  
 ii a. *hard, torque, term, burn*      b. *hammered*

The words in [i] all end in a vowel sound, while those in [ii] end in a vowel followed by just one consonant sound; note that the letter *e* at the end of the words in [ib] and of *torque* in [iia], and also that before the *d* in [iib] are ‘silent’ – i.e. there is no vowel in this position in the spoken form. In many of the non-rhotic accents such pairs of words as *mar* and *ma*, *floor* and *flaw*, or *torque* and *talk* are pronounced the same. A non-rhotic accent is thus one which lacks post-vocalic /r/.

Most speakers in the AmE family of dialects, by contrast, have a **rhotic accent**, where there is no such restriction on the distribution of /r/: all the words in [1] are pronounced with an /r/ sound after the (final) vowel, or (in the case of *stir* and *term*) with a rhoticised ('r-coloured') vowel sound, a coalescence of /r/ with the vowel.<sup>5</sup>

The English spelling system reflects the pronunciation of rhotic accents: in non-rhotic accents post-vocalic /r/ has been lost as a result of a historical change that took place after the writing system became standardised.

### ■ Linking and intrusive /r/

A further difference between non-rhotic and rhotic accents is seen in the pronunciation of such words and word sequences as those given in [2], where we use the symbol '·' to mark grammatical boundaries within a word (in these examples, between base and suffix):

- [2] i a. *marr·ing, sur·est, soar·ing*                      b. *the fear of death*  
     ii a. *saw·ing, thaw·ing*                                b. *the idea of death*

In non-rhotic accents the words in [ia] are all pronounced with /r/: the dropping of post-vocalic /r/ in the words *mar*, *sure*, *soar* does not apply here because the addition of a suffix beginning with a vowel makes the /r/ at the end of the base pre-vocalic. Similarly the word sequence [ib] is usually pronounced with an /r/ at the end of *fear* because the initial vowel of the next word makes it pre-vocalic.

The /r/ in pronunciations of [2i] in non-rhotic accents is called a **linking /r/**. Within a word, as in [ia], linking /r/ is obligatory; in word boundary position, as in [ib], the /r/ is optional though strongly preferred in most styles of speech. In [ii], where there is no *r* in the spelling, an /r/ pronounced at the end of the bases *saw·* and *thaw·* or of the word *idea* is called an **intrusive /r/**. Word-boundary intrusive /r/ in the pronunciation of sequences like [iib] is very common; word-internal intrusive /r/ in words like those in [iia] is much less common and quite widely disapproved of.

Rhotic accents do not have intrusive /r/ at all: they maintain a sharp distinction between [2i] and [ii], with /r/ appearing only in the former. And although they pronounce /r/ in the forms in [i], this is not linking /r/, since the bases *mar*, *sure*, *soar*, and *fear* have /r/ in these accents even when not followed by a vowel.

### 3.1.2 An accent-neutral phonological representation

Where we need to give pronunciations of words or larger expressions, it would be inconsistent with our goals to confine ourselves to one accent, but to attempt a complete listing of the pronunciations in each significant regional or other variety would be tedious. We therefore present here a unitary way of representing pronunciations for major BrE and AmE accents, whether rhotic or non-rhotic. For this purpose it is necessary to indicate more distinctions than would be needed in a system constructed for any one accent. In

<sup>5</sup>The correlation between the rhotic vs non-rhotic accent distinction and that between the BrE and AmE family of dialects is not perfect. Ireland, Scotland, the west of England, and some English-speaking Caribbean countries have rhotic accents and yet belong to the BrE family, and, conversely, there are various non-rhotic accents within the United States, including some working-class northeastern varieties and some upper-class southeastern varieties. The term 'rhotic' derives from the Greek name of the letter *r*.

particular, since it cannot be determined from the pronunciation in a non-rhotic accent where post-vocalic /r/ would occur in a rhotic one (for example, southern British English has /tɔ:k/ for both *torque* and *talk*), post-vocalic /r/ will have to be shown in some way even though it is not pronounced in the non-rhotic accents. Other differences have to be dealt with similarly.

The system we adopt is set out in [3], with illustrative examples in which the letter or letter sequence that symbolises the sound in question is underlined. Some notes on the system follow below.

[3] SHORT VOWELS

ɒ	<i>odd, lot, lost</i>	e	<i>get, fell, friend, endeavour</i>
æ	<i>gas, fat, pan</i>	i	<i>happy, pennies, maybe</i>
ʌ	<i>gut, much, done</i>	ɪ	<i>kit, build, women</i>
ə	<i>alone, potato, stringent, sofa</i>	ɪ	<i>wanted, luggage, buses</i>
ə <sup>r</sup>	<i>lunar, driver, actor</i>	ʊ	<i>look, good, put</i>

LONG VOWELS

ɑ:	<i>spa, calm, father</i>	ɔ:	<i>awe, dawn, caught, fall</i>
ɑ: <sup>r</sup>	<i>are, arm, spar</i>	ɔ: <sup>r</sup>	<i>or, corn, warn</i>
ɜ: <sup>r</sup>	<i>err, bird, work, fur</i>	u:	<i>ooze, blue, prune, brew, through</i>
i:	<i>eel, sea, fiend, dream, machine</i>		

DIPHTHONGS

aʊ	<i>owl, mouth, plough</i>	eə <sup>r</sup>	<i>air, bare, pear</i>
eɪ	<i>aim, day, eight, grey</i>	oʊ	<i>owe, go, dough, toe, goat</i>
aɪ	<i>I, right, fly, guy</i>	ɔɪ	<i>oil, boy</i>
ɪə	<i>idea</i>	ʊə <sup>r</sup>	<i>poor, sure, dour</i>
ɪə <sup>r</sup>	<i>ear, fear, pier, mere</i>		

TRIPHTHONGS

aɪə <sup>r</sup>	<i>ire, pyre, choir</i>	aʊə <sup>r</sup>	<i>our</i>
------------------	-------------------------	------------------	------------

CONSONANTS

b	<i>boy, sobbing</i>	ŋ	<i>sing, drink, dinghy</i>
d	<i>day, address</i>	θ	<i>thigh</i>
dʒ	<i>judge, giant, germ</i>	p	<i>pie</i>
ð	<i>this, although, bathe</i>	r	<i>rye, wrist</i>
f	<i>food, phonetics, if, off, rough</i>	s	<i>see, kiss, city, psychology</i>
g	<i>good, ghost, guide</i>	ʃ	<i>show, sure, charade, schmuck</i>
h	<i>hood</i>	t	<i>tall, pterodactyl</i>
j	<i>yes, fjord</i>	tʃ	<i>chin, watch</i>
k	<i>cat, chorus, kiss, brick, Iraqi</i>	v	<i>view, love, of</i>
l	<i>lie, all</i>	w	<i>wet</i>
m	<i>me, thumb, damn</i>	z	<i>zeal, peas</i>
n	<i>nigh, knife, gnaw, pneumatic</i>	ʒ	<i>measure, evasion, beige, rouge</i>

DIACRITICS

ɲ	syllabic /n/ (likewise for /l/, etc.)	'	stressed syllable ( <i>a'loof, 'sofa</i> )
---	---------------------------------------	---	--

## ■ Notes on the transcription system

### Post-vocalic /r/

This is represented by a superscript /<sup>r</sup>/. In rhotic AmE, it is pronounced as a separate /r/ consonant or coalesces with the preceding sound to give a rhotacised vowel. In non-rhotic BrE it is not pronounced at all – though, as we noted above, a word-final /<sup>r</sup>/ will typically be pronounced in connected speech as a pre-vocalic linking /r/ when followed by a word beginning with a vowel. Pre-vocalic /r/ corresponds to an *r* in the spelling. We do not include intrusive /r/ in our representations, since it is predictably present (between a low vowel or /ə/ and a following vowel) in those accents that have it.

### /ɪ/, /ə/, and /ɨ/

The unstressed vowel in the second syllable of *orange*, *wanted*, *wishes*, *lozenge*, etc., is a significant difficulty for an accent-neutral transcription. In BrE it is typically identical with the vowel of *kit*, which we represent as /ɪ/; in most AmE and some Australian varieties it is usually identical with the second vowel of *sofa*, /ə/. Many of its occurrences are in the inflectional endings; but there is one inflectional suffix in English that contains /ɪ/ in virtually all accents, namely *-ing*, and there are suffixes containing a vowel that is /ə/ in all accents (e.g. *-en* in *written*). Hence we need a third symbol for the vowel that varies between accents. We use /ɨ/. This has been used by American phonologists as a phonetic symbol for a vowel slightly less front than /ɪ/ and slightly higher than /ə/, so it is a good phonetic compromise, and visually suggests the /ɪ/ of those BrE accents that have a minimal contrast between *counted* /<sup>h</sup>kaʊntɪd/ and *countered* /<sup>h</sup>kaʊntəd/. It should be kept in mind, however, that it is used here not with an exact phonetic value but rather as a cover symbol for either /ɪ/ or /ə/ according to accent.

### /ɒ/ versus /ɑ/

For the vowel of *pot*, *rock*, *not*, etc., we use /ɒ/. Most varieties of AmE never have /ɒ/ phonetically in any context, so the American pronunciation can be derived simply by replacing our /ɒ/ by /ɑ/ everywhere. Hence there is no possibility of ambiguity.

### /oʊ/ versus /əʊ/

For the vowel of *grow*, *go*, *dough*, etc., we write /oʊ/, in which the ‘o’ makes the phonological representation closer to the spelling; for most BrE speakers /əʊ/ would be a phonetically more appropriate representation.

### /ɔː/ versus /ɑː/

BrE has distinct vowels in *caught* and *calm*: we represent them as /ɔː/ and /ɑː/ respectively. AmE standardly has the same vowel here, so for AmE the transcription /ɔː/ should be read as /ɑː/.

### /æ/ versus /ɑː/

Both BrE and AmE have distinct vowels in *fat*, /fæt/, and *calm*, /kɑːm/, but there are a considerable number of words where most BrE accents have /ɑː/ while AmE (but also some accents within the BrE family) has /æ/. Very few of these arise in our examples, however, so instead of introducing a third symbol we give separate BrE and AmE representations when necessary.

### /ʌ/ versus /ə/

The opposition between /ʌ/ and /ə/ is a weak one in that there are very few word-pairs kept distinct solely by this vowel quality difference. It is absent in many AmE accents – those in which *butt* and *but* are pronounced alike in all contexts, in which *just* has the same pronunciation whether it means “merely” or “righteous”, and in which *lust* always rhymes with *must* regardless of stress. We show the distinction between these vowels here (it is generally clear in BrE), but for many Americans both vowels could be written as /ə/.

### /ju:/ versus /u:/

In many words that have /ju:/ following an alveolar consonant in BrE, AmE has /u:/. Thus *new*, *tune*, *due* are /nju:/, /tju:n/, /dju:/ in BrE but usually /nu:/, /tu:n/, /du:/ in AmE. We write /ju:/ in these cases; for AmE, ignore the /j/.

### Intervocalic /t/

We ignore the AmE voicing of intervocalic /t/, contrasting *latter* as /lætə<sup>r</sup>/ and *ladder* as /lædə<sup>r</sup>/ with the medial consonants distinguished as in BrE accents.

## 3.2 Pronunciation and spelling

The relation between the sounds shown by our transcription and the ordinary English spelling of words is a complex one, and certain analytical concepts will help in keeping clear about the difference.

### ■ Symbols and letters

When we match up written and spoken forms we find that in the simplest cases one letter corresponds to one sound, or **phoneme**: *in* /ɪn/, *cat* /kæt/, *help* /help/, *stand* /stænd/, and so on. But very often the match is more complex. For example, in *teeth* the two-letter sequence *ee* corresponds to the single phoneme /i:/ and *th* to /θ/; in *plateau* the three-letter sequence *eau* corresponds to /oʊ/ (a diphthong, analysed phonologically as a single phoneme); in *through* the last four letters correspond to the phoneme /u:/.

We will use **symbol** as a technical term for a unit of writing that corresponds to a phoneme, and we will refer to those symbols consisting of more than one letter as **composite symbols**.<sup>6</sup> The letter *e* can form **discontinuous composite vowel symbols** with any of the letters *a*, *e*, *i*, *o*, *u*: *a . . . e* as in *pane*, *e . . . e* as in *dene*, *i . . . e* as in *bite*, *o . . . e* as in *rode*, and *u . . . e* as in *cute*.

### ■ Vowels and consonants

The categories **vowel** and **consonant** are defined in terms of speech. Vowels have unimpeded airflow through the throat and mouth, while consonants employ a significant constriction of the airflow somewhere in the oral tract (between the vocal cords and the lips). The terms can be applied to writing derivatively: a **vowel symbol** is a symbol representing a vowel sound, and a **consonant symbol** is a symbol representing a consonant sound. We will speak of a **vowel letter** or a **consonant letter** only in the case

<sup>6</sup>‘Digraph’ is widely used for a two-letter symbol and ‘trigraph’ is also found (though much less frequently) for a three-letter symbol, but there is no established term for a four-letter symbol, and no cover term for composite symbol.

of non-composite symbols: a single letter constituting a whole symbol may be called a vowel letter if it is a vowel symbol or a consonant letter if it is a consonant symbol. Thus *y* is a vowel letter in *fully* (representing /i/); it is a consonant letter in *yes* (it represents /j/); and in *boy* it is just part of a complex vowel symbol (representing /ɔɪ/). Similarly, *u* is a vowel letter in *fun* (/ʌ/), a consonant letter in *quick* (/w/), and part of a composite symbol in *mouth* (/aʊ/).<sup>7</sup>

It should be noted, however, that *r* counts as a consonant letter even in non-rhotic accents, as shown by the rule of final consonant letter doubling in inflected forms discussed in Ch. 18, §2.2.1: *map/mapping*, *bat/batting*, *trek/trekking*, *pin/pinning*, etc., are paralleled by *mar/marring*, with *r* doubling like other consonant letters. Similarly, the *e* of the suffix *-ed* counts as a vowel symbol even when no vowel is pronounced (e.g. it determines consonant doubling in forms like *sipped* [sɪpt] and *banned* [bænd]). In both cases, of course, the spelling corresponds more closely to an earlier stage of the language than to the contemporary language.

## 4 Theoretical framework

The primary goal of this grammar is to describe the grammatical principles of Present-day English rather than to defend or illustrate a theory of grammar. But the languages human beings use are too complex to be described except by means of a theory. In this section we clarify the relation between description and theory in this book, and outline some of our most important theoretical distinctions.

### 4.1 Description and theory

The problem with attempting to describe English without having a theory of grammar is that the language is too big to be described without bringing things together under generalisations, and without a theory there are no generalisations.

It does not take much reflection to see that there is no definite length limit to sentences in English. Sentences 100 words long, or longer, are commonly encountered (especially in writing, for written sentences are on average longer than spoken ones). And, given any sentence, it is always easy to see how it could have been made even longer: an adjective like *good* could be replaced by *very good*, or a verb like *exceed* could be supplied with a preceding adverb to make something like *dramatically exceed*, or a noun like *tree* could be replaced by *tall tree*, or the words *I think* could be added at the beginning of a whole declarative clause, or the words *and that's what I told the police* could be added at the end, and so on through an endless series of different ways in which almost any grammatical sentence of English could be lengthened without the result being something that is recognisably not English.

The importance of the fact that English sentences can be constructed to be as long as might be necessary to express some meaning is that it makes the sentences of English impossible to encapsulate in a list. The number of sentences that have been spoken or

<sup>7</sup>It will be clear, then, that we do not follow the traditional practice of simply dividing the alphabet into five vowels (*a, e, i, o, u*) and twenty-one consonants: we will see that the traditional classification does not provide a satisfactory basis for describing the spelling alternations in English morphology.

written so far is already astronomically vast, new ones are being produced every second around the world by hundreds of millions of people, and no matter what the information storage resources available, the problem is that there would be no way to decide where to end the list.

An alternative to listing sentences is therefore needed. To describe the sentences that belong to English we have to provide a general account of their structure that makes their form follow from **general** statements, not about particular sentences but about sentences of English quite generally. We need to bring together the principles that sentences all conform to, so that we can use those principles to appreciate the structure of new sentences as they are encountered, and see how new ones can be constructed. This means developing a **theory** of the ways in which sentences can be put together by combining words. This book is an attempt to summarise and illustrate as much as possible of what has so far been determined about the ways in which sentences can be constructed in English, and it presupposes a theory that classifies the words of the dictionary and specifies ways in which they are combined to form sentences.

We emphasise, however, that it is not the aim of this book to convince the reader of the merits of the theory for general linguistic description. Quite the reverse, in a sense: wherever it is possible to make a factual point overshadow a general theoretical point, we attempt to do that; whenever a theoretical digression would fail to illuminate further facts about English, we curtail the digression; if ever the facts at hand can be presented in a way that is neutral between competing theoretical frameworks, we try to present them that way.

However, a significant amount of space is devoted here to arguing carefully that the particular analysis we have decided to adopt, within the framework of theory we assume, is the right **analysis**. What we mean by that is that even someone with a different idea about how to design a theory of syntax would have to come to a conclusion tantamount to ours if they considered all the facts. It is necessary for us to provide arguments concerning specific grammatical analyses in this book because, although this grammar is descriptive like the great traditional grammars that have been published in the past, it is not traditional in accepting past claims and analyses.

We depart from the tradition of English grammar at many points, sometimes quite sharply. For example, in this book the reader will find nothing of ‘noun clauses’, ‘adjective clauses’, or ‘adverb clauses’, because that traditional distinction in subordinate clause classification does not divide things satisfactorily and we have abandoned it. The reader will likewise find nothing of the traditional distinction between *since* as a preposition (*I haven’t seen them since Easter*), *since* as an adverb (*I haven’t seen them since*), and *since* as a subordinating conjunction (*I haven’t seen them since they went overseas*), because we have concluded that this multiplication of categories for a single word with a single meaning makes no sense; we claim that *since* belongs to the same category (preposition) in all of its occurrences. On these and many other aspects of syntactic analysis we depart from traditional analyses (we draw attention to the major cases of this kind in Ch. 2). At such points we provide detailed arguments to convince the reader that we have broken with a mistaken tradition, and – we hope – made the correct decision about how to replace it.

The reader will therefore find much more discussion of grammatical concepts and much more syntactic argumentation than is usually found in grammars of English. It

is supplied, however, not to establish some wider theoretical point applying to other languages, but simply to persuade the reader that our description is sound. While the application of grammatical theories to the full range of human languages is an important matter within linguistics, it is not the purpose of this book to develop that point. Detailed technical or descriptive discussions that can be skipped by non-specialists without loss of continuity have been set off in smaller type with a shaded background.

## 4.2 Basic concepts in syntax

Three essential concepts figure in the theory we use to describe English syntax in this grammar. Each is very simple to grasp, but together they permit extremely broad and powerful theories to be constructed for indefinitely large collections of sentences. We express them tersely in [1].

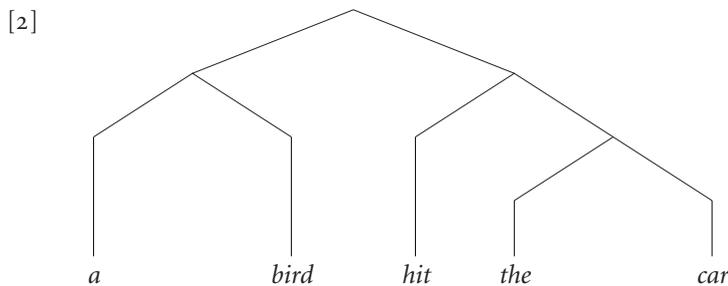
- [1] i Sentences have parts, which may themselves have parts.  
 ii The parts of sentences belong to a limited range of types.  
 iii The parts have specific roles or functions within the larger parts they belong to.

The idea that sentences have parts which themselves may have parts, i.e. that larger stretches of material in a sentence are made up by putting together smaller stretches, is the basis of ‘constituent structure’ analysis. The idea that the parts fall into a limited range of types that we can name and refer to when giving a grammatical description is the root of the concept of ‘syntactic categories’. And the idea that the parts also have specific roles or functions, or special slots that they fill in the larger parts they belong to, is the idea of ‘grammatical functions’. The next three subsections are devoted to explaining these three fundamental ideas.

### 4.2.1 Constituent structure

Sentences contain parts called **constituents**. Those constituents often have constituents themselves, and those are made up from still shorter constituents, and so on. This hierarchical composition of wholes from parts is called **constituent structure**.

Consider a simple one-clause sentence like *A bird hit the car*. It is divisible in the first instance into two parts, *a bird* (the subject) and *hit the car* (the predicate). The phrase *a bird* is itself made up of smaller parts, *a* and *bird*; so is *hit the car*, which we divide into *hit* and *the car*; and finally *the car* also has two parts, *the* and *car*. This structure can be represented as in [2].



Such representations of the constituent structure are called **trees** or tree-diagrams (though the trees are upside down, with the root at the top and the ends of the smallest

branches at the bottom). The words are the smallest constituents, and the points closer to the root where branches join identify the larger constituents. *A storm*, for example, is identified as a constituent because this word sequence can be traced via the branches to a single point in the tree; similarly with *the car* and *hit the car*. The sequence *bird hit*, on the other hand, is not a constituent, as there is no point in the tree that leads down branches to just these two words and no others.

The parts of the sentence shown at the first level down, *a bird* and *hit the car* are said to be the **immediate constituents** of the sentence; similarly, *hit* and *the car* are the immediate constituents of *hit the car*. The words are the **ultimate constituents** of the sentence.

The evidence that this is the correct analysis of the sentence comes from the whole of the rest of the grammar, all of which provides, by virtue of the coherence of the description it gives, the evidence that the lines of separation have been drawn in the right place. We can give an illustrative example of how other parts of the grammar can provide supportive evidence by considering where we can insert an adverb such as *apparently* (indicating that what the rest of the sentence asserts appears to be true). A rough account of where English grammar permits it to be positioned (at least in clauses as simple as our example) is that it can be anywhere in the clause it modifies, provided it does not interrupt a constituent. This is illustrated in [3], where the grammatical [a] examples conform to this rule, and the ungrammatical [b] ones do not:

- [3] i a. *Apparently a bird hit the car.*                      b. \**An apparently bird hit the car.*  
       ii a. *A bird apparently hit the car.*                     b. \**A bird hit apparently the car.*  
       iii a. *A bird hit the car, apparently.*                b. \**A bird hit the apparently car.*

The five words of our example sentence permit six different logically possible placements for *apparently* that are between words (before any of the five words, or after the last one), but only three are permissible. Breaking the sentence into constituents in exactly the way we have done, we are able to make a general statement about where an adverb like *apparently* (a ‘modal’ adverb) can be positioned in it: such an adverb must not interrupt a constituent of the clause. Hence [ib] above is disallowed because it would interrupt the constituent *a bird*; [iib] is disallowed because it would interrupt *hit the car*; and [iiib] is disallowed because it would interrupt *the car*. Inspecting the diagram in [2], we see that each of these uninterruptible sequences is a constituent smaller than the whole sentence.

The full support for a decision in grammatical description consists of confirmation from hundreds of mutually supportive pieces of evidence of many kinds, this being only one very simple example.

#### 4.2.2 Syntactic categories

Diagram [2] shows just the hierarchical part–whole relationships in the sentence. This is only the starting-point for a description, identifying the constituents that have to be described. The next step is to classify these constituents, to say what **syntactic category** they belong to. For words, these syntactic categories correspond to what are traditionally

called the ‘parts of speech’, and most of the categories for larger constituents are based on the ones for words. Where we need to refer to just the categories that have words as members, we will call them **lexical categories**.

### ■ Lexical categories

Any theory of syntax of the general sort we provide, and most types of dictionary, must include a list of the lexical categories or parts of speech assumed. For nearly all theories and nearly all dictionaries, **noun**, **verb**, **adjective**, and **adverb** will be among them, these being terms that have a history going back to the grammar of Classical Latin and Classical Greek some 2,000 years ago, but they are apparently applicable to almost all human languages. Our complete list is given, with some illustrations of membership, in [4]:

[4]	CATEGORY	LABEL	EXAMPLES
	i noun	N	<i>tree, pig, sugar, hatred, union, Picasso, London</i>
	ii verb	V	<i>do, fly, melt, think, damage, give, have, be, must</i>
	iii adjective	Adj	<i>good, nice, big, easy, ugly, helpful, reddish, fond</i>
	iv adverb	Adv	<i>obviously, easily, helpfully, frankly, soon, so, too</i>
	v preposition	Prep	<i>of, to, by, into, between, over, since, toward(s)</i>
	vi determinative	D	<i>the, this, that, a(n), some, all, every, each</i>
	vii subordinator		<i>that, for, to, whether, if</i>
	viii coordinator		<i>and, or, but, nor</i>
	ix interjection		<i>ah, damn, gosh, hey, oh, ooh, ouch, whoa, wow</i>

This scheme differs in several respects from the classification familiar from traditional grammar. Our determinatives are traditionally subsumed under the adjective category: they are said to be ‘limiting adjectives’ as distinct from the ‘descriptive adjectives’ illustrated in [4iii] – though some traditional grammars do recognise the articles *the* and *a(n)* as a distinct part of speech. We also take subordinators and coordinators to be distinct categories, not subclasses of the traditional conjunction category. Conversely, we regard pronouns as a subclass of nouns, not a distinct primary category. Our reasons for departing from the traditional analysis are given in the relevant chapters.

### ■ Phrasal categories

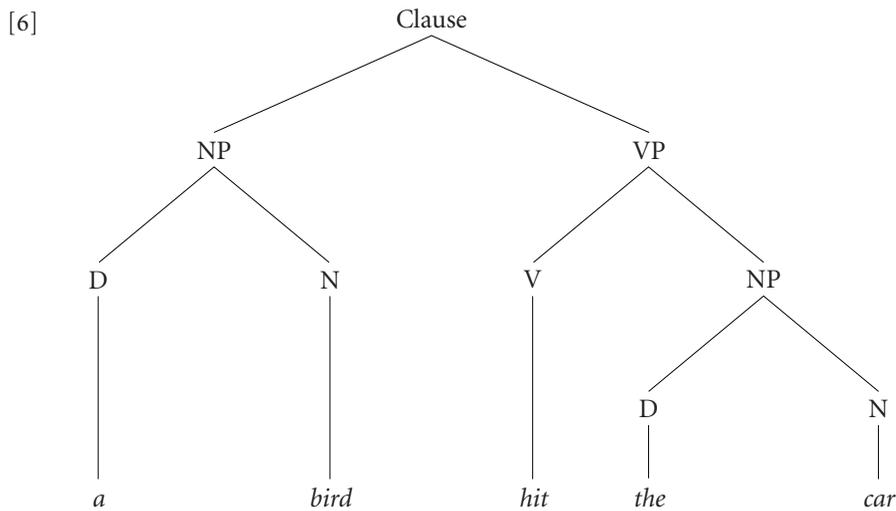
Constituents containing more than one word (more specifically, containing a central and most important word augmented by appropriate accompanying words that elaborate its contribution to the sentence) are called **phrases**, and are assigned to **phrasal categories**.<sup>8</sup> The lexical categories have corresponding phrase types that are in a sense expansions of them. A phrase consisting of a noun and the constituents that go with it most closely is a **nominal**; a nominal plus a determinative makes a **noun phrase**; a verb and its various complements makes up a **verb phrase**; a noun phrase and a verb phrase make up a **clause**; and so on. The full list of phrasal categories we employ in this book, together with our abbreviatory labels for them and an example phrase of each type, is given in [5].<sup>9</sup>

<sup>8</sup>There are circumstances in which phrases may consist of a single word: see the discussion of ‘singularity branching’ in §4.2.3.

<sup>9</sup>The term ‘sentence’ does not figure here. As will be explained more fully in Ch. 2, §1, a sentence in our terms is typically either a main clause or a coordination of main clauses.

[5]	CATEGORY	LABEL	EXAMPLE
i	clause	Clause	<i>she saw something in there</i>
ii	verb phrase	VP	<i>saw something in there</i>
iii	noun phrase	NP	<i>this clear case of dedication to duty</i>
iv	nominal	Nom	<i>clear case of dedication to duty</i>
v	adjective phrase	AdjP	<i>very eager for further news</i>
vi	adverb phrase	AdvP	<i>quite separately from this issue</i>
vii	preposition phrase	PP	<i>right out of the area</i>
viii	determinative phrase	DP	<i>almost every</i>

We can represent the structure of sentences in more detail than is done in a diagram like [2] if we show the category to which each constituent belongs, as in [6].



### 4.2.3 Grammatical constructions and functions

The third central theoretical idea we must introduce is that constituents always have particular roles to play in the **constructions**, the larger units, that they belong to. We call these roles **grammatical functions**. In our example sentence the phrases *a storm* and *the roof* belong to the same category, NP, but they have different functions, subject and object respectively. They belong to the same category because they are alike in their internal structure (both have a noun as the major element), but they have different functions because they stand in different relations to the verb. The opposite type of situation is illustrated in such a pair as:

- [7] a. His guilt was obvious.    b. That he was guilty was obvious.

Here the underlined constituents have the same function (subject) but belong to different categories (NP and clause respectively). They have the same function because they stand in the same relation to the predicate, and they belong to different categories because the first is centred on a noun (*guilt*) while the second is centred, ultimately, on a verb (*was*). We say that the subject is **realised** by an NP in [a], by a clause in [b].

## ■ Heads and dependents

There is a set of functions that to a large extent apply in the same way within all phrasal categories. The first division we make is that between the **head** and the various **dependents** that can combine with it.

The head, normally obligatory, plays the primary role in determining the **distribution** of the phrase, i.e. whereabouts in sentence structure it can occur. Note, then, that while *his guilt* and *that he was guilty* can both function as subject they differ in other aspects of their distribution – we can have, for example, *The news that he was guilty was devastating*, but not *\*The news his guilt was devastating* (we need a preposition: *The news of his guilt was devastating*), and this difference is attributable to the fact that the head of the former is a noun while the (ultimate) head of the latter is a verb.

Dependents, often optional, are syntactically subordinate elements. The term ‘dependent’ reflects the fact that in any given construction what kinds of dependent are permitted depends on the head. For example, *too* (with the sense “excessively”) can function as dependent to an adjective or adverb (*too careful, too carefully*), but not to a noun or verb (*\*their too extravagance, \*You shouldn’t too worry*). Similarly *sufficiently* can function as dependent to an adjective, adverb, or verb, but not to a noun (*sufficiently good, sufficiently often, practised sufficiently, \*sufficiently reason*).

### Predicate and predicator as special cases of the head function

Within this framework, what is traditionally called the **predicate** is a special case of the head function: the predicate is the head of a clause. Similarly, the term **predicator** is commonly used for the function of the verb itself, i.e. for the head of a verb phrase. We will retain the traditional terms, which indicate the characteristic semantic role of the element concerned, but it should be kept in mind that they are particular kinds of head.

## ■ Subtypes of dependent

Dependent is a very general function, and for many purposes we need to distinguish different subtypes of dependent according to their more specific relation to the head. At the first level of subdivision we distinguish **complements**, **modifiers**, and **determiners**, illustrated here in NP structure:

- [8] i *the photographs of their dog that they had brought with them* [complement]  
 ii *the photographs of their dog that they had brought with them* [modifier]  
 iii *the photographs of their dog that they had brought with them* [determiner]

In these examples, *of their dog* **complements** the head noun *photographs*; *that they had brought with them* **modifies** the head nominal noun *photographs of their dog*; and *the* **determines** the head nominal *photographs of their dog that they had brought with them*. At the next level we distinguish different kinds of complement, such as subject (*the photographs are excellent*), object (*He destroyed the photographs*), predicative (*these are excellent photographs*), and so on. A head element is said to **govern** its complements.

The determiner function is found only in the structure of the NP, whereas complements and modifiers occur quite generally. Note that the function ‘determiner’ is distinct from the lexical category ‘determinative’ (D). These need to be distinguished for the same

reason as we distinguish subject and NP. Thus although *this* functions as determiner in *this height*, it functions as modifier in the structure of an AdjP in examples like *She is about this tall*. Conversely, while the determiner function is realised by a determinative in *a doctor*, it is realised by a genitive NP in *my neighbour's doctor*.<sup>10</sup>

### ■ Non-headed constructions

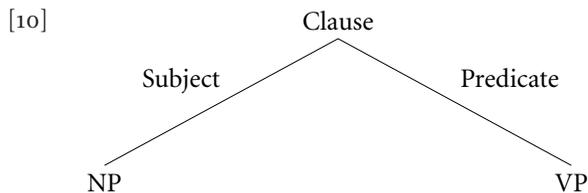
Although the functions of head and dependent apply to a very wide range of constructions, we must also allow for non-headed constructions, as in:

- [9] i *She bought [a hamburger, some chips and a glass of milk].* [coordination]  
 ii *A storm damaged – or so I'm told – the roof of their house.* [supplementation]

The underlined NPs in [i] are of equal syntactic status: we cannot say that one is head and the others dependents. Each of them has the same function within the bracketed construction, that of **coordinate**. In [ii] the underlined constituent is what we call a **supplement**: instead of being integrated into the constituent structure of the sentence as a dependent or coordinate, it is loosely attached, set off from the rest in speech by separate intonational phrasing and in writing by punctuation. Note that it interrupts the sentence at a point where a dependent could not occur, between the predicator and the object: compare [3iib] above.<sup>11</sup> These two types of non-headed construction are described in Ch. 15.

### ■ Diagrammatic representation of functions

Functions, we have said, are essentially relational concepts: to specify the function of a constituent is to say what its relation is to the construction containing it.<sup>12</sup> One way to capture this would be to write the name of the function on the line (branch) of the diagram joining the constituent to the construction. The first level in the structure of our model sentence might then look as in [10].

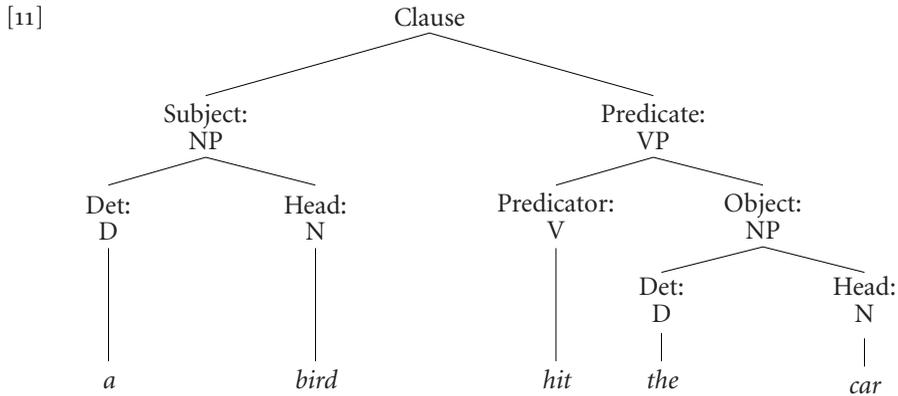


In more complex cases, though, diagram design becomes a problem, and we have found it preferable to present the functional labels separated from the category labels by a colon, and written above them in diagrams. In this format the analysis of our earlier example sentence looks as in [11].

<sup>10</sup>In other works ‘determiner’ is often used as a category term. The corresponding function is then sometimes called ‘specifier’, sometimes called ‘determinative’, and sometimes not clearly distinguished from the category term.

<sup>11</sup>It must be emphasised, therefore, that [3iib] was marked as ungrammatical with the understanding that *apparently* is integrated into the structure (as indicated by the absence of any punctuation). If *apparently* were set apart as a supplement, the sentence would not be ungrammatical – but it would be a different sentence from [3iib].

<sup>12</sup>‘Grammatical relation’ is indeed commonly used as an alternative term to ‘grammatical function’.



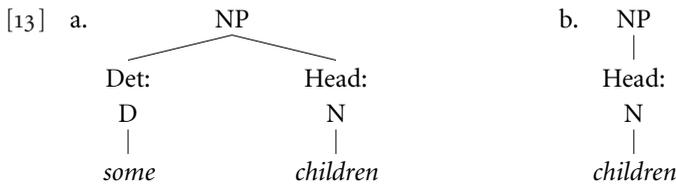
(Note that we use ‘Det’ as the abbreviation for the function ‘determiner’, and ‘D’ for the category ‘determinative’.)

### ■ Singular branching

We have said that dependents are often optional, and this implies that we can have a head on its own, without any dependents. Compare, for example:

- [12] i Some children were playing in the park.  
 ii Children were playing in the park.

The underlined expressions are NPs functioning as subject of the clause: *children* is the head, determined by *some* in [i], but standing alone in [ii]. The relevant parts of the structure are thus as in [13]. In [b] there is a single branch descending from the category



label NP, and this part of the tree-diagram is said to exhibit **singular branching**, in contrast to the **binary branching** of [a].

## 4.3 Morphology, inflectional and lexical

A grammar, we have said, is divided into two major components, syntax and morphology. This division follows from the special status of the word as a basic linguistic unit, with syntax dealing with the combination of words to make sentences, and morphology with the form of words themselves. In some respects the formation of words is comparable to the formation of larger units, but in others it is significantly different, and it is these differences that motivate dividing the grammar into two separate components.

### ■ Words, lexemes, and inflection

The term ‘word’, as used in traditional grammar, has two rather different senses. We can approach the difference by asking how many distinct words there are in, for example:

[14] *You are working hard, but your sister is working even harder.*

It is clear that the third and ninth words are not distinct: they are tokens (instances) of the same word. But what about *hard* and *harder*: are these the same word or different words? The answer depends on what you mean by ‘word’. In one sense they are obviously different: *harder* has a suffix that is missing from *hard*. This enables it to occur in constructions like that of *Your sister works harder than you*, where it could not be replaced by *hard*; and conversely *hard* could not be replaced by *harder* in *Your sister works very hard*. So from a syntactic point of view they are different words. But there’s another sense in which they are traditionally said to be ‘different forms of the same word’. The perspective this time is that of the dictionary, which would have just one entry, labelled *hard*. The same applies to *are* and *is* in [14]: syntactically these are different words, but lexically (i.e. as far as the dictionary is concerned) they are the same. In order to avoid possible misunderstanding we will restrict the term **word** to the syntactically-oriented sense, so that *hard* and *harder* are different words, and likewise *are* and *is*. For the more abstract, lexically-oriented sense we will use the term **lexeme**. *Hard* and *harder* are then forms of the same lexeme, as are *are* and *is*.

In many cases it makes no difference whether we take a syntactic or a lexical perspective. Lexemes such as *the* and *and* are **invariable**, i.e. there is only one word corresponding to each. Also invariable are lexemes like *efficiently*: although *more efficiently* is in some respects like *harder*, it is not a single word, but a sequence of two, and hence *efficiently* and *more efficiently* are not forms of a single lexeme. **Variable** lexemes, by contrast, are those which have two or more forms. Where we need to make clear that we are considering an item as a lexeme, not a word, we will represent it in bold italics. ***Hard***, for example, represents the lexeme which has *hard* and *harder* – and also *hardest* – as its forms.<sup>13</sup> Similarly *are* and *is*, along with *be*, *been*, *being*, etc., are forms of the lexeme ***be***. In example [14], then, we have two occurrences of the lexeme ***hard***, but only one of the word *hard*, and of course just one of the word *harder*. A variable lexeme is thus a word-sized lexical item considered in abstraction from grammatical properties that vary depending on the syntactic construction in which it appears.

The variation found in variable lexemes is known, more specifically, as **inflection**, and the various forms are called **inflectional forms** of the lexeme. For the most part, inflectional categories apply to large sets of lexemes. Almost all verb lexemes, for example, inflect for tense (e.g. preterite *took* vs present tense *take*), most nouns inflect for number (e.g. singular *dog* vs plural *dogs*), many adjectives one or two syllables in length inflect for grade (e.g. plain *old* vs comparative *older* vs superlative *oldest*). The inflectional contrast of nominative case vs accusative case (e.g. *we* vs *us*), however, applies to just a handful of pronoun lexemes.

<sup>13</sup> We minimise the use of bold type for lexemes, because in many cases it would simply distract. If we are simply listing adjective lexemes that can occur in the construction *They are difficult to please*, for example, we will generally list them as ‘*difficult, easy, hard, impossible, tough*’, etc., rather than ‘*difficult, easy, hard, impossible, tough*’, etc.; the fact that *easy* is inflectionally variable and *difficult* invariable has no relevance in that context.

### ■ Inflectional morphology and lexical word-formation

The distinction between words and lexemes provides the basis for the division of morphology into two branches: **inflectional morphology** and **lexical word-formation**.

Inflectional morphology deals with the inflectional forms of variable lexemes. It has something of the character of an appendix to the syntax, the major component of the grammar. Syntax tells us when a lexeme may or must carry a certain inflectional property, while inflectional morphology tells us what form it takes when it carries that inflectional property. For example, a rule of syntax stipulates that a verb in construction with the perfect auxiliary *have* must carry the past participle inflection (as in *They have killed it*, *She had rung the bell*), while inflectional morphology describes how the past participles of verbs are formed from the **lexical base**: *killed* is formed from the base *kill* by adding the suffix *·ed*, *rung* from *ring* by changing the vowel, and so on.

Lexical word-formation, by contrast, is related to the dictionary. It describes the processes by which new lexical bases are formed and the structure of complex lexical bases, those composed of more than one morphological element. The traditional term is simply ‘word-formation’: we add ‘lexical’ to exclude the formation of words by inflectional processes.

The three major processes involved in lexical word-formation are the following:

- [15] i COMPOUNDING: forming a new base by combining two bases  
 ii DERIVATION: forming a new base by adding an affix to an existing base  
 iii CONVERSION: forming a new base using the pronunciation/spelling of a base of related meaning in some other category

An example like *blackbird* illustrates compounding: it is formed by combining two smaller bases, *black* and *bird*. *Efficiently* illustrates derivation: an affix (the suffix *·ly*) is added to an adjective base (*efficient*) to form an adverb. Another example, this time not involving a change from one category to another, is the derivation of *inefficient* by adding the prefix *in-* to the same base. And conversion is illustrated by the underlined verb in *I managed to elbow my way to the front*. The base *elbow* is primarily a noun (having the singular form *elbow* and the plural form *elbows*) denoting a part of the body. The verb base *elbow* (the base of the lexeme whose forms are *elbow*, *elbows*, *elbowed*, *elbowing*) is formed from the noun by conversion – the shape of the noun is simply borrowed to make a verb of related meaning.

## 4.4 Defining grammatical concepts

A grammatical description of a language inevitably draws on a large repertoire of grammatical terms and concepts – **noun**, **verb**, **preterite**, **imperative**, **subject**, **object**, and countless more. A question arises concerning how these concepts are to be explained and defined.

### ■ Traditional grammar’s notional definitions

It is useful to begin by considering the kind of definition familiar from dictionaries and traditional school grammars, which are known as **notional definitions**, i.e. they are based on the meaning of the expressions being classified, not on their grammatical

properties. These are typical examples:

- |      |                 |   |                          |
|------|-----------------|---|--------------------------|
| [16] | i NOUN:         | the name of a person, place, or thing   | } [notional definitions] |
|      | ii PRETERITE:   | a tense expressing past action or state |                          |
|      | iii IMPERATIVE: | a clause expressing a command           |                          |

To determine whether a word is a noun, for example, one asks what it means or denotes; to determine the tense of a verb one asks in what time period it locates the action or state expressed by the verb; and so on.

Such definitions have long been criticised by linguists. Indeed, it takes only a moment or two's reflection to see that they do not provide satisfactory criteria for determining the correct classification of words or verb-forms or clauses. Take first the definition of preterite, and consider such examples as the following:

- |      |  |  |
|------|--|--|
| [17] | i a. <i>The finals <u>started</u> yesterday.</i> | b. <i>You said the finals <u>started</u> tomorrow.</i> |
|      | ii a. <i>I <u>gave</u> them his address.</i>     | b. <i>I regret <u>giving</u> them his address.</i>     |

In [i] we find *started* associated with past time in [a] but with future time in [b], as indicated by the temporal modifiers *yesterday* and *tomorrow* respectively. The *started* of [ia] thus satisfies the definition for preterite tense, while that of [ib] clearly does not. Nevertheless, everyone agrees that *started* in [ib] is a preterite form: this represents a different use of the same form as we have in [ia], not a different form, for the phenomenon is quite general, applying to all verbs, not just *start*. The opposite kind of problem arises in [ii]. Here the [a] and [b] versions are alike not in the form of the verb, but in the time of the associated event, which is located in the past. Both verbs therefore satisfy the definition of preterite tense, but while *gave* is certainly a preterite form, *giving* is not.

The notional definition thus gives the wrong results in both the [b] examples, excluding the *started* of [17ib], and including the *giving* of [iib]. If definitions are supposed to give necessary and sufficient conditions for belonging to some category, this one fails completely, for it gives neither: [ib] shows that past time reference is not necessary for a word to be a preterite verb form, and [iib] shows that it is not sufficient either. The problem is that the relation between the grammatical category of tense (form) and the semantic category of time (meaning) is highly complex, whereas the notional definition assumes the former can be defined directly in terms of the latter.

The same kind of problem arises with imperative clauses.<sup>14</sup> Compare:

- |      |                                     |  |
|------|-------------------------------------|--|
| [18] | i a. <i>Go to bed.</i>              | b. <i>Sleep well.</i>                      |
|      | ii a. <i>Please close the door.</i> | b. <i>Would you mind closing the door.</i> |

'Command', in the everyday sense of the term, is too narrow and specific for the meaning typically conveyed by imperatives: we will use the term 'directive' to cover commands, orders, requests, and other acts whose aim is to get the addressee to do something. With this modification, [ia] and [iia] both clearly satisfy the definition. But [ib] does not:

<sup>14</sup>Strictly speaking, the traditional category of imperative applies in the first instance to verb-forms rather than clauses. We take the view, however, that there are no imperative verb-forms in English, and hence consider the concept of imperative as it applies to clauses; the argument is in no way affected by this modification.

if I say this I am not telling you, or asking you, to sleep well, but expressing the hope or wish that you will. Yet grammatically it belongs with [ia] and [iia]: it is clearly an imperative clause. Conversely, [iib] conveys the same kind of meaning as [iia], but has a quite different grammatical structure: it is not imperative but interrogative. Again, then, satisfying the terms of the definition is not necessary for a clause to be imperative (as [ib] shows), nor is it sufficient (as [iib] shows). The relation between form and meaning here is too complex for one to be able to determine whether a clause is imperative or not simply on the basis of its meaning.

The traditional definition of noun is unsatisfactory for a somewhat different reason. The problem here is that the concept of ‘thing’ (or perhaps ‘name’) is too vague to provide a workable criterion. There are countless abstract nouns such as *absence*, *fact*, *flaw*, *idea*, *indeterminacy*, *lack*, *necessity*, etc., so ‘thing’ cannot be intended as equivalent to ‘physical object’; but we have no way of telling whether a word denotes (or is the name of) a thing unless we already know on independent, grammatical, grounds whether it is a noun. Take, for example:

- [19] i *I was annoyed at their rejection of my proposals.* [noun]  
 ii *I was annoyed that they rejected my proposals.* [verb]

These have essentially the same meaning, but *rejection* is a noun and *rejected* a verb. What enables us to tell that *rejection* but not *rejected* belongs to the category of noun is not that *rejection* denotes a thing while *rejected* does not, but that they figure in quite different grammatical constructions. Thus *rejection* contrasts with *rejections* as singular vs plural, whereas *rejected* contrasts with *reject* as preterite vs present tense. The transitive verb *rejected* takes a direct object (*my proposals*), but nouns do not take direct objects, so we need a prepositional complement in [i] (*of my proposals*). Similarly, *rejected* takes a nominative subject (*they*), whereas *rejection* takes a genitive like *their* or a determinative like *the*. And if we wanted to add some modification we would need an adjective in [i] (e.g. *their immediate rejection of my proposals*), but an adverb in [ii] (*that they immediately rejected my proposals*).

The problem with notional definitions is that they do not refer to the kinds of property that motivate the use in the grammar of the theoretical concepts being defined. The reason we need such concepts as noun, preterite, imperative clause in writing a grammar of English is that they enable us to make general statements about words, about the inflection of verbs, about the structure of clauses. Lexemes fall into a number of major categories on the basis of their inflection, the kinds of dependent they take and the function in larger constructions of the phrases they head: noun belongs in this system of lexeme categories. Verbs have a variety of inflectional forms, and the preterite is one of these. Clauses show structural contrasts on one dimension according to the presence or absence of a subject, its position relative to the verb, and the inflectional form of the verb, so that we have contrasts between such sets as (a) *You are punctual*, (b) *Are you punctual?*, (c) *Be punctual*: ‘imperative clause’ is one of the terms in this system of clausal constructions.

A satisfactory definition or explanation of concepts like noun, preterite, and imperative clause must therefore identify the grammatical properties that distinguish them from the concepts with which they contrast. The discussion of *rejection* and *rejected* in [19] illustrated some of the major ways in which nouns differ from verbs. As for the

preterite, it is distinguished in part by its form (in regular verbs it is marked by the suffix *·ed*, though this also marks the past participle), in part by its distribution (like the present tense, but unlike other forms, a preterite form can occur as the verb of a declarative main clause: *Kim gave it away*, but not, for example, *\*Kim given it away*), in part by its lack of agreement with the subject (with the single exception of the verb **be**), and so on. Imperative clauses differ from declaratives and interrogatives in the form of the verb **be** (*Be punctual* vs *You are punctual*), the optionality of a 2nd person subject (*you* is omissible in *You be punctual*, but not in *You are punctual*), the formation of the negative (compare *Don't be punctual*, formed with auxiliary **do**, and *You aren't punctual*, with no **do**), and so on.

In this grammar we will be at pains, therefore, to specify the distinctive grammatical properties of the concepts we introduce. This is not to suggest that we are not interested in the meaning, but rather to say that we need to distinguish between grammatical concepts and semantic ones; indeed, making such a distinction is a prerequisite for describing the relation between them.

#### ■ General and language-particular definitions

In criticising the traditional notional definitions, we assumed that they were intended to enable us to determine what expressions in English belong to the categories concerned. It must be emphasised, however, that most of the terms that figure in a grammatical description of English are not unique to English but appear in the grammars of other languages too – in some cases, in the grammars of all languages. There are therefore two issues to be considered in defining or explaining such terms. At one level there is the issue of what grammatical properties distinguish one category from another in English. We call this the **language-particular** level. This is the level we have been concerned with so far. A language-particular definition will enable us to decide which expressions in the language concerned belong to the category. At another level there is the issue of what principled basis we have for using the same grammatical terms in the grammars of different languages, given that the language-particular distinctive properties will vary from language to language. We call this the **general** level. The fact, for example, that the negative imperative *Don't be punctual* requires auxiliary **do** while the negative declarative *You aren't punctual* does not is clearly a specific fact about English: it belongs in the language-particular definition of imperative clause for English, but not in a general definition.

It might then be suggested that the traditional notional definitions should be construed as applying at the general rather than the language-particular level. Certainly they are not intended to apply uniquely to English. But at the same time there can be no doubt that as they are presented in school textbooks, for example, they purport to be language-particular definitions: the student is meant to be able to apply them to decide whether a given word in English is a noun, whether a verb is in the preterite, whether a clause is imperative. In effect, the traditional definitions aim to work at both levels simultaneously, and our objection is that the levels need to be distinguished, and approached differently. At the language-particular level, as we have argued, it is necessary to focus on form: to specify the grammatical features that distinguish expressions which belong to the category from those that do not. At the general level it is quite legitimate to invoke meaning: languages serve to express meaning, and it

is rare to find grammatical distinctions that have no correlation at all with semantic distinctions.

We need to make it clear when giving a general definition that it is to apply at the general level, not the language-particular. And we need to acknowledge that the correlation between grammatical form and meaning is typically complex rather than one-to-one. The general definitions we propose for the categories discussed above are as follows:

- [20] i NOUN: a grammatically distinct category of lexemes of which the morphologically most elementary members characteristically denote types of physical objects (such as human beings, other biological organisms, and natural or artificial inanimate objects)
- ii PRETERITE: a grammatically distinct inflectional form of the verb whose primary use is to locate the situation in past time (relative to the time of utterance)
- iii IMPERATIVE: a grammatically distinct clause construction whose members are characteristically used to issue directives

The move to an avowedly general definition, together with the reference to characteristic use of the most elementary members, enables us to avoid the vagueness of the term ‘thing’ (or ‘name’) in [16]. The fact that such lexemes as *rejection*, *arrival*, *idea* do not denote physical objects is not a problem for a definition at this level. By virtue of the distinctive grammatical properties specified in the language-particular definition, these lexemes belong to the same category as *girl*, *boy*, *daffodil*, *window*, etc., and this category as a whole satisfies the general definition of noun because it contains lexemes like these last examples that do denote physical objects. Note that the abstract nouns *rejection* and *arrival* are morphologically derived from lexemes of another category (verb); morphologically elementary nouns, such as *girl*, *boy*, etc., characteristically denote kinds of physical object.

Definition [20ii] allows for the fact that verb inflections often have more than one use. In [17ia] (*The finals started yesterday*), we have the past time use. In [17ib] (*You said the finals started tomorrow*) the preterite form *started* is within a subordinate clause functioning as complement to *said*: this is a case of what is traditionally called indirect reported speech. Your actual words will have been, say, *The finals start tomorrow*, but present tense *start* is shifted into the preterite *started* in my report. Another use of the preterite is seen in *I wish the finals started tomorrow*, where it indicates counterfactuality: we understand that the finals do not start tomorrow. Of these three uses, it is the one that indicates past time that is primary. The others are found only in special contexts, such as the complement of a preterite verb of reporting or the complement of *wish*. This verb-form therefore qualifies for the label **preterite**.

Definition [20iii] likewise overcomes the problems we noted in [16iii]. The language-particular criteria assign *Sleep well* to the same category as *Go to bed* and *Please open the door*, and since most clauses with this form are normally used as directives we call the category imperative clause. *Would you mind closing the door* is excluded from the category at the language-particular level: it does not have the distinctive grammatical form of imperative clauses in English.

### The grammatical distinctiveness requirement in general definitions

It will be noted that the general definitions in [20] all impose a condition of grammatical distinctiveness. This requirement means that the general term being defined will be applicable in the grammar of a given language only if it can be given a distinct language-particular definition in that language.

A significant weakness of traditional grammars of English is that they incorporate a number of categories that in fact have no place in a grammar of Present-day English, although they are perfectly valid for Latin (and in some cases older stages of English). A simple example is provided by the dative case inflection. A traditional dictionary or schoolbook definition is given in [21i], while our proposed revision is given in [ii]:

- [21] i DATIVE: the case of nouns, etc., expressing the indirect object or recipient  
 ii DATIVE: a grammatically distinct case characteristically used to mark the indirect object

Definition [i] suggests that in *He gave Caesar a sword*, for example, *Caesar* is in the dative case, as it is in indirect object function and expresses the semantic role of recipient. And that indeed is the analysis found in many traditional grammars and school textbooks (especially older ones). But Present-day English has no dative case. In the Latin counterpart of the above sentence *Caesar* has a different form (*Caesari*) from the one it has when functioning as subject (*Caesar*) or direct object (*Caesarem*), so the distinctiveness condition of definition [ii] is satisfied for Latin. In English it is not satisfied: the form is simply *Caesar* whether the function is subject, direct object, or indirect object. There is no noun, not even a pronoun, with a distinct inflectional form for the indirect object, and hence no basis at all for including dative among the inflectional categories of the English noun.<sup>15</sup>

## 5 Semantics, pragmatics, and meaning relations

Few grammars even attempt to describe the ways in which sentences are formed without making reference along the way to meaning and how sentences express it. After all, few would take it to be controversial that a human language such as English is in some sense a system for framing thoughts and making meaningful messages expressible, and this would make it a natural supposition that meaning and grammar would be to some extent intertwined. This grammar, while not attempting a full and detailed semantic description of the language (which would be an unrealistically large and difficult enterprise), touches on the topic of meaning frequently. But as we will explain, we do not treat meaning as a unitary phenomenon.

### ■ The semantics/pragmatics distinction

We treat the analysis of meaning as divisible in the first instance into two major domains. The first deals with the sense conventionally assigned to sentences independently of the contexts in which they might be uttered. This is the domain called **semantics**. The second

<sup>15</sup>Our definition omits the reference to recipients in the traditional definition because this will appear in the definition of indirect object – a grammatically distinct subtype of object characteristically expressing the recipient.



throne would allow [ia] to continue to assert a true proposition, but [ib] would assert a false proposition under those circumstances. The sentences accordingly have different truth **conditions**: circumstances could obtain under which one would express a truth and the other a falsehood. Similarly, though perhaps less obviously, in [ii]. For [iia] to be true, it is sufficient for the committee to feel broadly favourable to my plan, but for [iib] to be true it is necessary that they actually took some action to give my plan the go-ahead signal. The conditions under which the first would be true are not quite the same as those under which the second would be true, so the meanings differ.

### ■ Entailments

One way of describing truth conditions is in terms of **entailments**. An entailment is defined as follows (the definitions in this chapter use ‘ $\equiv$ ’ to symbolise the relation ‘is by definition equivalent to’):

[3]  $X$  entails  $Y \equiv$  If  $X$  is true, then it follows necessarily that  $Y$  is true too.

In the first instance, entailment is a relation between propositions, since it is propositions, strictly speaking, that have truth values. But we can apply the concept derivatively to sentences, as illustrated in:

[4]	i	<i>Kim broke the vase.</i>	[entails [ii]]
	ii	<i>The vase broke.</i>	[entailed by [i]]
	iii	<i>Kim moved the vase.</i>	[does not entail [ii]]

If the proposition asserted by [i] in any context is true, then the proposition asserted by [ii] in that same context must also be true. The first proposition entails the second, and sentence [i] entails sentence [ii]. If  $X$  entails  $Y$ , then it is inconsistent to assert  $X$  and deny  $Y$ . It is inconsistent, for example, to say <sup>#</sup>*Kim broke the vase but the vase didn't break* (the ‘#’ symbol indicates that what follows is grammatical but semantically or pragmatically anomalous). In the case of [iii] and [ii] there is no such inconsistency: *Kim moved the vase but the vase didn't break*. And [iii] of course does not entail [ii]: it is perfectly possible for [iii] to be true and [ii] false.

We can state entailments in a variety of equivalent ways: we can say that *Kim broke the vase* entails that the vase broke, or that it entails “The vase broke”, or that it entails *The vase broke*. Whichever mode of presentation we adopt, it follows from the definition given in [3] that if  $X$  entails  $Y$  then  $X$  cannot be true unless  $Y$  is true. And that is to say that  $Y$  is a condition for the truth of  $X$ . So to give the entailments of a sentence is to give its truth conditions.

### ■ Closed and open propositions

A refinement of our notion of proposition is called for in discussing certain constructions. What we have described so far as propositions could be described more precisely as **closed propositions**. They are closed in the sense of not leaving anything available to be filled in: a proposition like “Sandy showed me that at the office last week” identifies what was done, who did the showing, what was shown, where it happened, and when this occurred. There are also **open propositions**, which have a place left open. Consider the meaning of *What did Sandy show you at the office last week?*: it could be represented informally as “Sandy showed you  $x$  at the office last week”, where  $x$  is a placeholder, or **variable**, for a piece of information not supplied. The point of open interrogative sentences like *What*



is some relevant contrast related to the second coordinate – perhaps one would have expected her to use a return flight and she is acting counter to that expectation, or it might be that although she will be going up there at air travel speed she will have much more time for reading on the slow return trip, and so on. The precise nature of the contrast is not made explicit, but the use of *but* rather than the neutral coordinator *and* indicates that the two parts are being presented as involving some sort of contrast. As we have said, this extra meaning contributed by the choice of *but* rather than *and* is not part of the propositional meaning: it would not be legitimate for you to respond to [ib] by saying, *That's false, though I concede that she is flying up there and taking the train back.*

Similarly with [6ii], except that here the two sentences differ not in the choice of one word rather than another, but in the presence or absence of a word, namely *even*. *Even* conveys that it is somehow noteworthy that the property of having agreed that his behaviour was outrageous applies to Max: it is less expected that Max should have agreed than that the others who agreed should have done so. Again, this is not part of the propositional meaning. The truth conditions of [iia–iib] are the same: there is no context where one could be true and the other false. But it is intuitively clear that the sentences do not have exactly the same meaning.

The same applies in [6iii]. *Bloody* serves in some rather vague way to express anger or ill will towards sales tax reporting regulations, or towards the idea of having to work out sales taxes, or something of the sort. But the anger or ill will is not expressed as part of the propositional meaning: the truth conditions for [iiib] are exactly the same as those for [iiia].

We will handle the non-propositional meaning conveyed by items such as *but*, *even*, and *bloody* in these examples in terms of the concept of **conventional implicature**. In uttering [6ib], I indicate, or **implicate**, that there is some kind of contrast between her taking the train back and flying up there, but I do not actually state that there is. And analogously for the others. Unlike entailments, conventional implicatures are not restricted to sentences that are characteristically used to make statements. *Is she flying up there but taking the train back?*, *Did even Max agree that his behaviour had been outrageous?* and *Have you ever had to do a bloody sales tax report?* carry the above implicatures even though they do not themselves have truth conditions.

### 5.3 Pragmatics and conversational implicatures

Pragmatics is concerned not with the meaning of sentences as units of the language system but with the interpretation of utterances in context. Utterances in context are often interpreted in ways that cannot be accounted for simply in terms of the meaning of the sentence uttered.

Let us again illustrate the point by means of a few representative examples:

- [7] i *Do you think I could borrow five dollars from you?*  
 ii *If you agree to look after my horses after I die, I'll leave you my whole estate.*  
 iii *Some of the audience left the room before the first speaker had finished.*

Imagine that Sue and Jill are at the cash register in a cafeteria buying sandwiches. Jill has \$20 in her hand. Sue finds she only has a few cents in her purse, and utters [i]. As far as the literal meaning of the sentence is concerned, this is a question as to whether or not

Jill thinks Sue could borrow five dollars from her. It has two possible answers: “Yes” (i.e. “I do think you could”) and “No” (i.e. “I don’t think you could”). But for Jill to respond *Yes, I do* would seem strange and uncooperative in this context. It would force Sue to be more direct: *Well, lend it to me then, right now, because I can’t afford to pay for this sandwich.*

What would normally be expected of Jill would be to act on the basis of the following reasoning. We both have to pay for our sandwiches. Sue has reached the cash register and, after finding her purse almost empty, is asking whether in my opinion it would be possible for me to extend a \$5 loan. Sue can see that I have \$20, and sandwiches only cost about \$5, so I could obviously afford it. Sue must see that the answer to the question is “yes”. Why am I being asked for my opinion about my financial status? What is the point of this question? The only reasonable conclusion is that Sue actually wants me to advance such a loan, right now.

The message “Please lend me \$5” is thus indirectly conveyed by a question that does not itself actually express it. A cooperative addressee will understand the speaker’s intention immediately, without consciously going through the process of reasoning just sketched. But for the student of language it is important to see: (a) that “Please lend me \$5” is not the semantic meaning of sentence [7i], but the pragmatic meaning of an utterance of [i] in a certain range of contexts; (b) that the pragmatic interpretation can be derived in a systematic way from the interaction between the sentence meaning and the context.

Semantics is thus concerned with the meaning that is directly expressed, or encoded, in sentences, while pragmatics deals with the principles that account for the way utterances are actually interpreted in context. A central principle in pragmatics, which drives a great deal of the utterance interpretation process, is that the addressee of an utterance will expect it to be **relevant**, and will normally interpret it on that basis.

This principle of relevance was very evident in our first example: the relevance of Sue’s question was that she needed Jill to lend her the money. It is equally important in deriving the pragmatic interpretation of [7ii]. This sentence does not actually make the statement that you won’t get the estate if you don’t agree to look after my horses: that is not part of the sentence meaning. A proposition of the type “if P then Q” does not require “P” to be true in order for “Q” to be true.<sup>16</sup> We therefore need an explanation for this fact: anyone who is told *If you agree to look after my horses after I die then I’ll leave you my whole estate* will always assume that the bequest will not be forthcoming without the agreement to look after the horses. Why? Because otherwise it would not have been relevant to mention the horses. If that part of the sentence had some relevance, it must be as a necessary condition for getting the bequest, and we normally try to find an interpretation for an utterance that makes everything in it relevant. The semantics of the sentence does not tell us that the horse care will be a precondition for the bequest, but the pragmatics of interpreting the utterance certainly does.

<sup>16</sup>If this is not obvious, consider the sentence *If a house collapses directly on me I will die*. This does not entail that provided no house falls on me I will be immortal. Eventually I will die anyway. Or consider *If you need some more milk there’s plenty in the fridge*. This does not state that there is plenty of milk in the fridge only if you need some. If there is milk in there, it will be there whether you need it or not. A sentence meaning “if P then Q” will often strongly suggest “if not P then not Q”, but that is not part of the meaning.

Consider, finally, example [7iii], as uttered, say, in the context of my giving you an account of a weekend seminar I recently attended. You will infer that not all of the audience left the room before the first speaker had finished. But again that is not part of the meaning of the sentence. *Some* does not **mean** “not all”. The “not all” interpretation can be accounted for by pragmatic principles. I am describing an event at which I was present, so I presumably know whether or not all of the audience left before the first speaker had finished. Suppose I know that all of them left. Then I would surely be expected to say so: such a mass walkout would be much more worth mentioning than one where only part of the audience left. So the natural assumption is that I said *some* rather than *all* because it would not have been true to say *all*: what other reason could I have for making the weaker statement?

Compare this with the case where you ask *Have all the questionnaires been returned?* and I reply *I don't know: some have, but I can't say whether they all have*. If *some* meant “not all” this would be incoherent, but clearly it is not. This time my reason for saying *some* rather than *all* is not that it would be false to say *all*, but merely that I do not have enough knowledge or evidence to justify saying *all*.

We will again invoke the concept of implicature in describing the above interpretations of utterances of [7i–iii], but we will classify them more specifically as **conversational implicatures**. We will say, for example, that an utterance of [7iii] in the context described conversationally implicates “Not all of the audience left before the first speaker had finished”.

#### ■ Relation between entailment and the two kinds of implicature

The differences between entailment, conventional implicature, and conversational implicature are summarised in [8].

[8]	ENTAILMENT	semantic	truth-conditional
	CONVENTIONAL IMPLICATURE	semantic	non-truth-conditional
	CONVERSATIONAL IMPLICATURE	pragmatic	non-truth-conditional

Implicatures are distinguished from entailments in that they are not truth conditions; hence they are not restricted to sentences that can be used to make statements. The two types of implicature are distinguished according to whether they are part of the conventional meaning of sentences or derive from the interaction between the sentence meaning and the context of utterance by means of general principles of conversational cooperation. In this book we will be much more concerned with conversational implicatures than with conventional ones, as they play a larger part in the interpretation of discourse; we will take them to represent the default case, therefore, and when the term implicature is used without qualification it is intended to be understood in the conversational sense in the absence of indications to the contrary. The verb corresponding to ‘implicature’ is **implicate**; in addition, we will use the term **convey** in a way which is neutral between entail and (conventionally or conversationally) implicate.

Conversational implicatures are not part of sentence meaning at all. They are suggested to the hearer by the combination of the sentence meaning and the context, but they are not part of what is **said**. Nevertheless, many of them are of very general application, so that we can say that such-and-such an implicature will normally accompany the utterance of a given sentence unless special factors exclude that possibility. In such cases

it is convenient to talk about the sentence normally implicating something – e.g. that [7iii] normally implicates that not all of the audience left before the first speaker had finished. This is to be understood as a shorthand way of saying that an utterance of the sentence in a normal context would carry that implicature in the absence of factors which exclude it. We will therefore apply the term to sentences in the following sense:

- [9]  $X$  normally **con conversationally implicates**  $Y \equiv X$  does not entail  $Y$  but in saying  $X$  the speaker makes an implicit commitment to the truth of  $Y$  in the absence of indications to the contrary.

When such ‘indications to the contrary’ are present, we will say that the implicature is **cancelled**. Take, for example:

- [10] *Some if not all of the delegates had been questioned by the police.*

Without the underlined sequence, *some* would again trigger a “not all” implicature – that not all of the delegates had been questioned by the police. This implicature, however, is inconsistent with *if not all*, which explicitly allows for the possibility that all of the delegates had been questioned. The implicature is therefore cancelled, i.e. is here not part of the interpretation. A context where the request-to-borrow implicature of [7i] could be cancelled might be one where I’m concerned with the legality of borrowing: perhaps I’m the treasurer of some institution and am uncertain whether I am permitted to go into debt.

The possibility of cancellation is an essential feature of conversational implicatures. If something conveyed by an utterance were an invariable component of the interpretation of the sentence, whatever the context, it would be part of the sentence meaning, either a conventional implicature or an entailment. Some conversational implicatures, however, are very **strong** in the sense that it is not easy to imagine them being cancelled – and these run the risk of being mistaken for components of sentence meaning. But it is important to make the distinction. It would be impossible, for example, to give a satisfactory account of quantification in the noun phrase if the “not all” component in the interpretation of *some* were not recognised as merely a conversational implicature.

## 5.4 Pragmatic presupposition

Finally, we consider the relation of **presupposition**, exemplified in:

- [11] i *She has stopped trying to secure her son’s release.*  
 ii *She hasn’t stopped trying to secure her son’s release.*  
 iii *Has she stopped trying to secure her son’s release?*  
 iv *She formerly tried to secure her son’s release.* } [all presuppose [iv]]

Presupposition has to do with informational status. The information contained in a presupposition is backgrounded, taken for granted, presented as something that is not currently at issue. In [11] all of [i–iii] presuppose that she formerly tried to secure her son’s release: what is at issue is not whether she tried to secure his release in the past but whether she is doing so now.

This example brings out an important property of presupposition, namely that it is generally unaffected by negation or questioning. When a sentence is negated, the negation characteristically applies to that part of the content that is presented as being at

issue. If she in fact never tried to secure her son's release, [ii] is strictly speaking true, but it would normally be a very inefficient or misleading way of conveying that information. A simpler, more direct and more explicit way of doing so would be to say *She never tried to secure her son's release*. The fact that I didn't say this but said [ii] instead will lead you to infer that the negation applies to the stopping, so that [ii] implicates that she is still trying. Similarly with questioning. If I didn't know, and wanted to find out, whether she formerly tried to secure her son's release, I would be expected to ask *Did she try to secure her son's release?* If I ask [iii] instead, the natural inference will be that I am trying to find out about the present state of affairs.

The kind of reasoning just described is similar in kind to that invoked in discussing conversational implicatures, reflecting the fact that both phenomena are pragmatic.<sup>17</sup> Like conversational implicature, presupposition applies in the first instance to utterances, but we can apply it derivatively to sentences with the same 'normally' qualification as before:

- [12] *X normally presupposes Y*  $\equiv$  in saying *X* the speaker, in the absence of indications to the contrary, takes the truth of *Y* for granted, i.e. presents it as something that is not at issue.

Again, then, we allow that in special circumstances a presupposition may be cancelled. Consider, for example, the following exchange:

- [13] *A: Have you stopped using bold face for emphasis?*  
*B: No I haven't (stopped using bold face for emphasis); I've always used small caps.*

A's question presupposes that B formerly used bold face for emphasis. But suppose it turns out that A was mistaken in believing this. B answers the question with a negative, and since this reflects the form of the question it too would normally presuppose that B formerly used bold face for emphasis. But in the context given here that presupposition is cancelled.

The presupposition associated with the verb *stop* coincides with an entailment when *X* is positive and declarative, as in [11i], but with a conversational implicature when *X* is negative or interrogative, as in [11ii–iii]. You cannot stop doing something that you have never done before, so [11i] cannot be true unless [11iv] is true. This gives the latter the status of an entailment. But it is not an entailment of the negative [11ii], as evident from the example in [13]. Nevertheless, if I say [11ii] I will normally be taken to have implicitly committed myself to [11iv], and the latter therefore counts as a conversational implicature. Likewise with the interrogative [11iii], which does not have entailments.

This represents the most usual pattern for presuppositions. For the most part they are entailed if *X* is positive and asserted to be true, and otherwise they are conversationally implicated. But this is not a necessary feature of presuppositions: we will see that they do not always follow this pattern.

<sup>17</sup> An alternative view is that presupposition is a logical or semantic concept. On one version of this account, a presupposition is a proposition that must be true if the presupposing proposition (or the sentence expressing it) is to be either true or false. In the case of [11], for example, in a context where [iv] was false, where she had never tried to secure her son's release, [i–ii] would be neither true nor false: they would simply lack a truth value (or would take a third truth value distinct from both truth and falsity). We do not adopt that concept of presupposition here, and take the view that if a proposition is not true, then it is false.

