1 Grammar

1.1 Overview

In broad terms, this book is concerned with aspects of grammar. Grammar is traditionally subdivided into two different but interrelated areas of study – **morphology** and **syntax**. Morphology is the study of how words are formed out of smaller units (called **morphemes**), and so addresses questions such as 'What are the component morphemes of a word like *antidisestablishmentarianism*, and what is the nature of the morphological operations by which they are combined together to form the overall word?' Syntax is the study of the way in which phrases and sentences are structured out of words, and so addresses questions like 'What is the structure of a sentence like *What's the president doing?* and what is the nature of the grammatical operations by which its component words are combined together to form the overall sentence structure?' In this chapter, we begin (in §1.2) by taking a brief look at the approach to the study of syntax taken in **traditional grammar**: this also provides an opportunity to introduce some useful grammatical terminology. In the remainder of the chapter, we look at the approach to syntax adopted within the theory of **Universal Grammar** developed by Chomsky.

1.2 Traditional grammar: Categories and functions

Within traditional grammar, the syntax of a language is described in terms of a taxonomy (i.e. classificatory list) of the range of different types of syntactic structures found in the language. The central assumption underpinning syntactic analysis in traditional grammar is that phrases and sentences are built up of a series of **constituents** (i.e. syntactic units), each of which belongs to a specific **grammatical category** and serves a specific **grammatical function**. Given this assumption, the task of the linguist in analysing the syntactic structure of any given type of sentence is to identify each of the constituents in the sentence, and (for each constituent) to say what category it belongs to and what function it serves. For example, in relation to the syntax of a simple sentence like:

(1) Students protested

it would traditionally be said that the sentence consists of two constituents (the word *students* and the word *protested*), that each of these constituents belongs to a specific grammatical category (*students* being a plural **noun** and *protested* a past tense **verb**) and that each serves a specific grammatical function (*students* being the **subject** of the sentence, and *protested* being the **predicate**). The overall sentence *Students protested* has the categorial status of a **clause** which is **finite** in nature (by virtue of denoting an event taking place at a specific time), and has the semantic function of expressing a **proposition** which is **declarative** in **force** (in that it is used to make a statement rather than e.g. ask a question). Accordingly, a traditional grammar of English would tell us that the simplest type of finite declarative clause found in English is a sentence like (1), in which a nominal subject is followed by a verbal predicate. Let's briefly look at some of the terminology used here.

In traditional grammar, words are assigned to grammatical categories (called **parts of speech**) on the basis of their **semantic** properties (i.e. meaning), **morphological** properties (i.e. the range of different forms they have) and **syntactic** properties (i.e. word-order properties relating to the positions they can occupy within sentences): a set of words which belong to the same category thus have a number of semantic, morphological and syntactic properties in common. There are traditionally said to be two different types of word, namely **content words/contentives** (= words which have substantive lexical content) on the one hand, and **function words/functors** (= words which essentially serve to mark grammatical properties) on the other. The differences between the two can be illustrated by comparing a contentive like *car* with a functor like *they*. A noun like *car* has substantive lexical content in that it denotes an object which typically has four wheels and an engine, and it would be easy enough to draw a picture of a typical *car*; by contrast, a pronoun such as *they* has no descriptive content (e.g. you can't draw a picture of *they*), but rather is a functor which simply marks grammatical (more specifically, person, number and case) properties in that it is a third person plural nominative pronoun. Because they have lexical semantic content, content words often (though not always) have antonyms (i.e. 'opposites') – e.g. the adjective *tall* has the antonym *short*, the verb *increase* has the antonym *decrease*, and the preposition *inside* has the antonym *outside*: by contrast, a typical function word like e.g. the pronoun *me* has no obvious antonym. Corresponding to these two different types of (content and function) word are two different kinds of grammatical category – namely **lexical/substantive categories** (= categories whose members are content words) on the one hand, and **functional categories** (= categories whose members are function words) on the other.

Let's begin by looking at the main **lexical/substantive categories** found in English – namely **noun, verb, adjective, adverb** and **preposition** (conventionally abbreviated to N, V, A, ADV and P in order to save space). **Nouns** (= N) are traditionally said to have the semantic property that they denote entities: so, *bottle* is a noun (since it denotes a type of object used to contain liquids),

water is a noun (since it denotes a type of liquid) and *John* is a noun (since it denotes a specific person). There are a number of distinct subtypes of noun: for example, a noun like *chair* is a **count noun** in that it can be counted (cf. *one chair, two chairs* . . .), whereas a noun like *furniture* is a **mass noun** in that it denotes an uncountable mass (hence the ungrammaticality of **one furniture,* **two furnitures* – a prefixed star/asterisk being used to indicate that an expression is ungrammatical). Likewise, a distinction is traditionally drawn between a **common noun** like *boy* (which can be modified by a determiner like *the* – as in *The boy is lying*) and a **proper noun** like *Andrew* (which cannot be used in the same way in English, as we see from the ungrammaticality of **The Andrew is lying*). Count nouns generally have the morphological property that they have two different forms: a **singular** form (like *horse* in *one horse*) used to denote a single entity, and a **plural** form (like *horses* in *two horses*) used to denote more than one entity. Common nouns have the syntactic property that only (an appropriate kind of) noun can be used to end a sentence such as *They have no* . . . In place of the dots here we could insert a singular count noun like *car,* or a plural count noun like *friends* or a mass noun like *money,* but not other types of word (e.g. not *see* or *slowly* or *up*, as these are not nouns).

A second lexical/substantive category is that of **verb** (= V). These are traditionally said to have the semantic property that they denote actions or events: so, *eat*, *sing*, *pull* and *resign* are all (action-denoting) verbs. From a syntactic point of view, verbs have the property that only an appropriate kind of verb (in its uninflected **infinitive** form) can be used to complete a sentence such as *They/It can* . . . So, words like *stay*, *leave*, *hide*, *die*, *starve* and *cry* are all verbs and hence can be used in place of the dots here (but words like *apple*, *under*, *pink* and *if* aren't). From a morphological point of view, regular verbs like *cry* in English have the property that they have four distinct forms: e.g. alongside the **bare** (i.e. uninflected) **form** *cry* we find the **present tense** form *cries*, the **past tense/perfect participle/passive participle** form *cried* and the **progressive participle** form *crying*. (See the Glossary of terminology at the end of this book if you are not familiar with these terms.)

A third lexical/substantive category is that of **adjective** (= A). These are traditionally said to have the semantic property of denoting states or attributes (cf. *ill*, *happy*, *tired*, *conscientious*, *red*, *cruel, old* etc.). They have the syntactic property that they can occur after *be* to complete a sentence like *They may be* . . . (as with *They may be tired/ill/happy* etc.), and the further syntactic property that (if they denote a **gradable** property which can exist in varying degrees) they can be modified by a degree word like *very/rather/somewhat* (cf. *She is very happy*). Many (but not all) adjectives have the morphological property that they have **comparative** forms ending in *-er* and **superlative** forms ending in *-est* (cf. *big/bigger/biggest*).

A fourth lexical/substantive category is that of **adverb** (= ADV). These often have the semantic property that they denote the manner in which an action is performed (as with *well* in *She sings well*). Regular adverbs have the morphological

property that they are formed from adjectives by the addition of the suffix *-ly* (so that corresponding to the adjective *sad* we have the adverb *sadly*). A syntactic property of adverbs is that an adverb (like e.g. *badly*) is the only kind of word which could be used to end sentences such as *She behaved . . .*, *He treats her . . .* or *He worded the statement . . .*

The fifth and final lexical/substantive category found in English is that of **preposition** (= P). Many of these have the semantic property of marking location (cf. *in/on/off/inside/outside/under/above/below*). They have the syntactic property that a preposition (with the appropriate kind of meaning) can be modified by *right* in the sense of 'completely', or by *straight* in the sense of 'directly' (as with the preposition *down* in *He fell right down the stairs* and the preposition *to* in *He went straight to bed*). Prepositions have the morphological property that they are invariable/uninflected forms (e.g. the preposition *off* has no past tense form **offed*, no superlative form **offest* and so on).

In addition to the five lexical/substantive categories identified above, English also has a number of functional categories. One such functional category is that of **determiner** (= D) – a category whose members are traditionally said to include the definite article *the* and the demonstrative determiners *this/that/these/those*. They are called determiners because they have the semantic property that they determine specific semantic properties of the noun expression that they introduce, marking it as a definite referring expression: for example, an expression like *the car* in a sentence such as *Shall we take the car?* is a definite referring expression in the sense that it refers to a definite (specific) car which is assumed to be familiar to the hearer/addressee. A related class of words are those which belong to the functional category **quantifier** (= Q), denoting expressions of quantity, such as *some/all/no/any/each/every/most/much/many*. (We shall also take the indefinite article *a* to be a quantifier – one which quantifies over a single entity.)

A further type of functional category found in English is that of **pronoun** (= PRN). Pronouns are items which are said to 'stand in place of' (the meaning of the prefix *pro-*) or 'refer back to' noun expressions. However, there are reasons to think that there are a number of different types of pronoun found in English and other languages. For example, in sentences such as *John has a red car and Jim has a blue one*, the word *one* is traditionally said to be a pronoun because it has no lexical semantic content of its own, but rather takes its content from its **antecedent** (i.e. *one* refers back to the noun *car* and so *one* is interpreted as having the same meaning as *car*). However, from a morphological perspective, the pronoun *one* behaves like a regular count noun in that it has a plural form ending in *-s* (as in *I'll take the green apples if you haven't got any red ones*). So, more accurately, we could say that *one* is an N-pronoun (or pronominal noun). By contrast, in a sentence like *Many miners were rescued, but some died*, the word *some* seems to function as a Q-pronoun (i.e. a pronominal quantifier). And in a sentence like *These apples are ripe, but those aren't*, the word *those* seems to be a D-pronoun (i.e. a pronominal determiner). Indeed, some linguists have argued that so-called **personal pronouns** like *I/me/we/us/you/he/him/she/her/it/they/them* are also

D-pronouns: the rationale for this is that some such pronouns can be used as determiners which modify a following noun (as in *We republicans don't trust you democrats*, where *we* could be argued to be a determiner modifying the noun *republicans*, and *you* could be seen as a determiner modifying the noun *democrats*). While, as noted here, pronouns can be argued to belong to a number of distinct types of category, in order to simplify discussion I shall simply refer to them as belonging to the category PRN throughout this book. (Because there are a number of different types of pronoun, some linguists prefer to refer to them by using the more general term **proform**.)

Another type of functional category found in English is that of **auxiliary (verb)**. They have the semantic property of marking grammatical properties such as **tense**, **aspect**, **voice** or **mood** (see the Glossary of terminology at the end of the book if you are not sure what these terms mean). Auxiliaries have the syntactic property that (unlike lexical/main verbs) they can be inverted with their subject in questions (so that corresponding to a statement like *It is raining* we have the question *Is it raining?* where the auxiliary *is* has moved in front of the subject *it* and is said to have been *inverted*). The items italicised in (2) below (in the use illustrated there) are traditionally categorised as auxiliaries taking a [bracketed] complement containing a bold-printed verb:

(2) (a) He *has*/*had* [**gone**]
(b) She *is*/*was* [**staying** at home]
(c) They *are*/*were* [**taken** away for questioning]
(d) He really *does*/*did* [**say** a lot]
(e) You *can*/*could* [**help** us]
(f) They *may*/*might* [**come** back]
(g) He *will*/*would* [**get** upset]
(h) I *shall*/*should* [**return**]

In the uses illustrated here, *have/be* in (2a,b) are (**perfect/progressive**) **aspect** auxiliaries, *be* in (2c) is a (**passive**) **voice** auxiliary, *do* in (2d) is an **expletive** or **dummy** auxiliary (i.e. one with no intrinsic lexical semantic content), and *can*/*could*/*may*/*might*/*will*/*would*/*shall*/*should* in (2e–h) are **modal** auxiliaries. What auxiliaries in sentences like those above have in common is the fact that they inflect for present/past tense. Hence, in work in syntax over the past ten years or so, they have been said to belong to the category T (= tense-marked auxiliary).

An interesting word which has been argued to be related to tense-marking auxiliaries in work over the past thirty years or so is the infinitive particle *to*, in sentences such as:

(3) They are now expecting the president *to* be impeached tomorrow

In a sentence like (3), infinitival *to* seems to have future time-reference (in that the act of impeachment will take place at some time in the future), and this is why we can use the word *tomorrow* in the *to*-clause. In this respect, infinitival *to* seems

to have much the same function as the auxiliary *will* in *They are now expecting that the president will be impeached tomorrow*, suggesting that infinitival *to* is an **infinitival tense marker**, and so belongs to the same category T as present/past tense auxiliaries such as *is/was*. The difference between auxiliaries and infinitival *to* is that most auxiliaries overtly inflect for present/past tense (though this is not true of the invariable auxiliaries *must* and *ought*), whereas infinitival *to* is invariable in form. We can thus say that an auxiliary like *will* is a finite T constituent, whereas infinitival *to* is a nonfinite T.

The last type of functional category which we will look at is a kind of word (like each of the words italicised in the examples below) which is traditionally termed a (subordinating) **conjunction:**

(4) (a) I think [*that* you may be right]
(b) I doubt [*if* you can help me]
(c) I'm anxious [*for* you to receive the best treatment possible]

Each of the bracketed clauses in (4) is a complement clause, in that it is the complement of the word immediately preceding it (*think*/*doubt*/*anxious*); for this reason, the italicised word which introduces each clause is known in work since the 1960s as a **complementiser** (= C), and this is the terminology which will be adopted throughout this book. Complementisers are functors in the sense that they encode particular sets of grammatical properties. For example, complementisers encode (non)finiteness by virtue of the fact that they are intrinsically finite or nonfinite. More specifically, the complementisers *that* and *if* are inherently finite in the sense that they can only be used to introduce a finite clause (i.e. a clause containing a present or past tense auxiliary or verb, like the present tense auxiliaries *may* and *can* in 4a and 4b); by contrast, *for* is an inherently infinitival complementiser, and so can be used to introduce a clause containing infinitival *to* (as in 4c). Moreover, *that* introduces a **declarative** clause (i.e. one which has the **force** of a statement), *if* introduces an **interrogative** clause (i.e. one which has the force of a question) and *for* introduces an **irrealis** clause (i.e. one relating to a hypothetical event which hasn't yet taken place and may or may not take place at some stage in the future). Hence, we can say *that* is a finite declarative complementiser, *if* is a finite interrogative complementiser and *for* is an infinitival irrealis complementiser.

Using the set of syntactic categories outlined above, we can employ the traditional **labelled bracketing** technique to *categorise* words (i.e. assign them to grammatical categories) in a way which describes how they are being used in a particular sentence. Using this technique, the words in sentence (5a) below can be categorised as in (5b):

(5) (a) The president is clearly feeling angry that Congress has refused to negotiate with him
(b) [$_{D}$ The] [$_{N}$ president] [$_{T}$ is] [$_{ADV}$ clearly] [$_{V}$ feeling] [$_{A}$ angry] [$_{C}$ that] [$_{N}$ Congress] [$_{T}$ has] [$_{V}$ refused] [$_{T}$ to] [$_{V}$ negotiate] [$_{P}$ with] [$_{PRN}$ him]

The labelled bracketing in (5b) tells us that *the* is a D/determiner, *president* a N/noun, *is* a T/present tense auxiliary, *clearly* an ADV/adverb, *feeling* a V/verb, *angry* an A/adjective, *that* a C/complementiser, *Congress* a N/noun, *has* a T/present tense auxiliary, *refused* a V/verb, *to* a T/infinitival tense particle, *negotiate* a V/verb, *with* a P/preposition and *him* a PRN/pronoun.

The discussion of grammatical categories presented above is merely a brief sketch: however, it suffices to illustrate the point that when traditional grammarians analyse the syntax of sentences, they begin by assigning each of the words in the sentence to a grammatical category which describes how it is being used in the sentence concerned. Grammatical differences between individual words belonging to the same category are traditionally described in terms of sets of **grammatical features**, and these features (by convention) are enclosed in square brackets. For example, both *she* and *us* are pronouns, but they differ in that *she* is a **third person** pronoun which is **feminine** in **gender**, **singular** in **number** and **nominative** in **case**, whereas *us* is a first person pronoun which is **plural** in number and **accusative** in case. Accordingly, we can describe the differences between these two pronouns by saying that the pronoun *she* carries the features [third-person, singular-number, feminine-gender, nominative-case], whereas *us* carries the features [first-person, plural-number, accusative-case].

As noted at the beginning of this section, traditional grammarians are also concerned to describe the **grammatical functions** which words and other expressions fulfil within the sentences containing them. We can illustrate this point in terms of the following set of sentences:

(6) (a) *John* smokes
(b) *The president* smokes
(c) *The president of Utopia* smokes
(d) *The former president of the island paradise of Utopia* smokes

Sentence (6a) comprises the noun *John* which serves the function of being the **subject** of the sentence (and denotes the person performing the act of smoking), and the verb *smokes* which serves the function of being the **predicate** of the sentence (and describes the act being performed). In (6a), the subject is the single noun *John*; but as the examples in (6b,c,d) show, the subject of a sentence can also be an (italicised) phrase like *the president*, or *the president of Utopia* or *the former president of the island paradise of Utopia*.

Now consider the following set of sentences:

(7) (a) John smokes *cigars*
(b) John smokes *Cuban cigars*
(c) John smokes *Cuban cigars imported from Havana*
(d) John smokes *a specific brand of Cuban cigars imported by a friend of his from Havana*

Sentence (7a) comprises the **subject** *John*, the **predicate** *smokes* and the **complement** (or **direct object**) *cigars*. (The complement *cigars* describes the entity on

which the act of smoking is being performed; as this example illustrates, subjects normally precede the verb with which they are associated in English, whereas complements typically follow the verb.) The complement in (7a) is the single noun *cigars*; but a complement can also be a **phrase**: in (7b), the complement of *smokes* is the phrase *Cuban cigars*; in (7c) the complement is the phrase *Cuban cigars imported from Havana*; and in (7d) the complement is the phrase *a specific brand of Cuban cigars imported by a friend of his from Havana*. A verb which has a noun or pronoun expression as its direct object complement is traditionally said to be **transitive**.

From a semantic perspective, subjects and complements share in common the fact that they generally represent entities directly involved in the particular action or event described by the predicate: to use the relevant semantic terminology, we can say that subjects and complements are **arguments** of the predicate with which they are associated. Predicates may have one or more arguments, as we see from sentences such as (8) below, where each of the bracketed nouns is a different argument of the italicised predicate:

(8) (a) [John] *resigned*
(b) [John] *felt* [remorse]
(c) [John] *sent* [Mary] [flowers]

A predicate like *resign* in (8a) which has a single argument is said to function as a **one-place predicate** (in the relevant use); one like *feel* in (8b) which has two arguments is a **two-place predicate**; and one like *send* in (8c) which has three arguments is a **three-place predicate**.

In addition to predicates and arguments, sentences can also contain **adjuncts**, as we can illustrate in relation to (9) below:

(9) (a) The president smokes a cigar *after dinner*
(b) The president smokes a cigar *in his office*

In both sentences in (9), *smokes* functions as a two-place predicate whose two arguments are its subject *the president* and its complement *a cigar*. But what is the function of the phrase *after dinner* which also occurs in (9a)? Since *after dinner* isn't one of the entities directly involved in the act of smoking (i.e. it isn't consuming or being consumed), it isn't an argument of the predicate *smoke*. On the contrary, *after dinner* simply serves to provide additional information about the time when the smoking activity takes place. In much the same way, the italicised expression *in his office* in (9b) provides additional information about the location of the smoking activity. An expression which serves to provide (optional) additional information about the time or place (or manner, or purpose etc.) of an activity or event is said to serve as an **adjunct**. So, *after dinner* and *in his office* in (9a,b) are both **adjuncts**.

So far, all the sentences we have looked at in (6–9) have been **simple sentences** which contain a single **clause**. However, alongside these we also find **complex sentences** which contain more than one clause, like (10) below:

(10) Mary knows John smokes

If we take the traditional definition of a clause as a predication structure (more precisely, a structure containing a predicate which has a subject, and which may or may not also contain one or more complements and adjuncts), it follows that since there are two predicates (*knows* and *smokes*) in (10), there are correspondingly two clauses – the *smokes* clause on the one hand, and the *knows* clause on the other. The *smokes* clause comprises the subject *John* and the predicate *smokes*; the *knows* clause comprises the subject *Mary*, the predicate *knows* and the complement *John smokes*. So, the complement of *knows* here is itself a clause – namely the clause *John smokes*. More precisely, the *smokes* clause is a **complement clause** (because it serves as the complement of *knows*), while the *knows* clause is the **main clause** (or **principal clause** or **independent clause** or **root clause**). The overall sentence (10) *Mary knows John smokes* is a **complex sentence** because it contains more than one clause. In much the same way, (11) below is also a complex sentence:

(11) The press clearly think the president deliberately lied to Congress

Once again, it comprises two clauses – one containing the predicate *think*, the other containing the predicate *lie*. The main clause comprises the subject *the press*, the adjunct *clearly*, the predicate *think* and the complement clause *the president deliberately lied to Congress*. The complement clause in turn comprises the subject *the president*, the adjunct *deliberately*, the predicate *lie* and the complement *to Congress*.

As was implicit in our earlier classification of (1) as a **finite** clause, traditional grammars draw a distinction between **finite** and **nonfinite** clauses. In this connection, consider the contrast between the italicised clauses below (all of which function as the complement of an underlined adjective or verb):

(12) (a) She was glad *that he apologised*
(b) She demanded *that he apologise*
(c) I can't imagine *him apologising*
(d) It would be sensible *for him to apologise*
(e) It's important to know *when to apologise*

The italicised clauses in (12a,b) are finite, and it is characteristic of finite clauses in English that they contain an (auxiliary or main) verb marked for tense/mood, and can have a nominative pronoun like *he* as their subject. In (12a), the verb *apologised* is finite by virtue of being inflected for past tense and **indicative mood**, and by virtue of having a nominative subject (*he*); in (12b), the verb *apologise* is finite by virtue of being inflected for **subjunctive mood** (and perhaps present tense, though this is far from clear), and by virtue of having a nominative subject (*he*). A clause containing a verb in the indicative mood denotes a real (or **realis**, to use the relevant grammatical term) event or state occurring at

a specific point in time; a subjunctive clause by contrast denotes a hypothetical or unreal (= **irrealis**) event or state which has not yet occurred and which may never occur. In contrast to the italicised clauses in (12a,b), the clauses italicised in (12c–e) are nonfinite, in that they contain no verb marked for tense or mood, and do not allow a nominative subject. For example, the verb *apologising* in (12c) is nonfinite because it is a tenseless and moodless **gerund** form, and has an **accusative** subject *him*. Likewise, the verb *apologise* in (12d,e) is a tenseless and moodless **infinitive** form (as we see from the fact that it follows the infinitive particle *to*), and has an accusative subject *him* in (12d), and a 'silent' (implicit) subject in (12e). (Excluded from our discussion here are gerund structures with genitive subjects like the italicised in 'I can't stand *his perpetual(ly) whining about syntax*', since these are more nominal than clausal in nature.)

As the examples in (12) illustrate, whether or not a clause is finite in turn determines the kind of subject it can have, in that finite clauses can have a **nominative** pronoun like *he* as their subject, but nonfinite clauses cannot. Accordingly, one way of telling whether a particular clause is finite or not is to see whether it can have a nominative pronoun (like *I/we/he/she/they*) as its subject. In this connection, consider whether the italicised clauses in the dialogues in (13a,b) below are finite or nonfinite:

(13) (a) SPEAKER A: I know you cheat on me
SPEAKER B: OK, I admit it. *I cheat on you*. But not with any of your friends
(b) SPEAKER A: I know you cheat on me
SPEAKER B: *Me cheat on you?* No way! I never would!

The fact that the italicised clause in speaker B's reply in (13a) has the nominative subject *I* suggests that it is finite, and hence that the verb *cheat* (as used in the italicised sentence in 13a) is a first person singular present tense form. By contrast, the fact that the italicised clause in speaker B's reply (13b) has the accusative subject *me* suggests that it is nonfinite, and that the verb *cheat* (as used in the italicised sentence in 13b) is an infinitive form (and indeed this is clear from sentences like *Me be a cheat? No way!* where we find the infinitive form *be*).

In addition to being finite or nonfinite, each clause within a sentence has a specific **force**. In this connection, consider the following simple (single-clause) sentences:

(14) (a) He went home
(b) Are you feeling OK?
(c) You be quiet!
(d) What a great idea that is!

A sentence like (14a) is traditionally said to be **declarative** in force, in that it is used to make a statement. (14b) is **interrogative** in force in that it is used to ask a question. (14c) is **imperative** in force, by virtue of being used to issue an order