

1 How the Language Is Made Up

A language is not just a conglomeration of words. It is a complex system of mental organisation, with structures and systems flowing one into the other. Each section of the organisation depends upon other sections, and is in part determined by them. Grammar provides the infrastructure, and lexical words fill grammatical slots. The resulting product enables people to communicate – to work together, to compose songs and stories, to construct scientific argumentation, to express their emotions, and many things besides.

It is a little like a physical organisation. There will be sectors for purchasing raw materials, for manufacturing, for packaging, for sales, for advertising, and so on. Each sector is linked to the others. The salespeople provide feedback to packaging on what appeals most to customers. Advertising extols the purity of the raw materials used. Within the organisation there are people – filling roles, doing jobs, checking, and planning.

It would not be informative to say that the company consists essentially of a set of people, with an organisation being built around them. It is the organisation that is prior, and suitable people are chosen to make it work. Similarly for language. It would not be informative to state that a story, say, consists of a collection of words with grammar being wrapped around them. The story has a structure. Connected paragraphs are made up of coordinated sentences, each with predicate and subject; the latter will have a central element (the head) and an optional set of modifiers. Lexical words are chosen to fill slots in the structure (a bit like people in the manufacturing company).

As an illustration, there are two basic grammatical frames in which adjectives may occur: after the copula verb *be* (this is called ‘copula complement’ function) and modifying a noun within a noun phrase. *Happy* and *content* are adjectives with similar meanings. But to know how to use them properly, one must be aware of how they operate within the overall grammatical organisation of the language.

Happy is used in both frames:

The manager was happy with the decision reached
She was a happy manager

In contrast, *content* occurs as copula complement:

The manager was content with the decision reached

But *content* may not be used to modify a noun; that is, it is not acceptable to say ♦*She was a content manager*.

Now consider two more adjectives with similar meanings, *lone* and *isolated*. Both may modify a noun:

There was a lone house on the plain

There was an isolated house on the plain

But only *isolated* may function as copula complement:

The house on the plain was isolated

That is, one does not say ♦*The house on the plain was lone*.

It is not enough to state that a word is an adjective, with a certain meaning. The grammatical framework of the language must first be established in order to understand the ways in which each adjective may be used. And similarly for every other kind of word.

We can now examine a pair of verbs, *give* and *donate*. They involve three participants: Donor, Gift, and Recipient. *Give*, with a general meaning, occurs in two grammatical frames:

The philanthropist_{DONOR} gave £500,000_{GIFT} to [the hospital]_{RECIPIENT}
 The philanthropist_{DONOR} gave [the hospital]_{RECIPIENT} £500,000_{GIFT}

Donate refers to a particular type of giving, where the Recipient is a worthy cause. It can occur in the first construction available for *give*:

The philanthropist_{DONOR} donated £500,000_{GIFT} to [the hospital]_{RECIPIENT}

but not in the second one. That is, ♦*The philanthropist_{DONOR} donated [the hospital]_{RECIPIENT} £500,000_{GIFT}* is not an acceptable sentence for most speakers.

A dictionary should explain the meaning of each word, and the grammatical contexts in which it may be used. These are linked. For example, *donate* focuses on the magnitude of the Gift (others will be giving to the same Recipient) and, in view of this, the noun phrase referring to the Gift must be in direct object function, immediately after the verb.

A language consists of two independent but interlocking parts: lexicon (or vocabulary) and grammar. Plus phonology, which codes meanings into sounds – these are uttered by a speaker, heard by a listener, and decoded by them back into meanings. And orthography, the representation in writing of phonology.

A language makes up a single large system, each part of which only has significance with respect to the whole. Taking it one step further, a language

does not exist in a vacuum. It is a social phenomenon, reflecting the way of life of its community of speakers, mirroring the world in which they live. This determines the grammatical possibilities, the assemblage of lexical words and their meanings, and how these knit together. (Exemplified by the way in which *donate* can be used.)

In the remainder of this chapter we characterise grammar and lexicon (illustrating grammar with discussion of comparative constructions and complement clauses), then discuss phonological and orthographic representations, and the multi-faceted nature of ‘word’. The final section deals with two approaches to language description – vocabulary prior, or an integration of grammar and vocabulary.

Grammar and Lexicon

The grammatical template of a language deals with types of main clause. An intransitive clause has a verb plus a subject noun phrase (NP), as in *The boy yawned*. A transitive clause has a verb plus a subject NP and an object NP, as in *The pensioner has eaten lunch*. In various circumstances, a transitive clause may omit the object NP; one could just say *The pensioner has eaten*. However, a subject NP can never be omitted.

There are also subordinate clauses, one important type being relative clauses. We can start with two simple clauses:

- (1) I saw the dog
- (2) The dog chased the cat

The second of these can be incorporated into the object NP of the first:

- (3) I saw [the dog [which chased the cat]_{RELATIVE.CLAUSE}]_{OBJECT.NP}

Here, the relative clause helps to specify which dog it was that I saw.

The relative clause is introduced by *which*. One of the arguments of the relative clause must be identical to the head (or central element) of the NP in which it occurs. In (3) it is the subject of the relative clause, *the dog*. Alternatively, it could be the object, *the cat*, as in:

- (4) I saw [the cat [(which) the dog chased]_{RELATIVE.CLAUSE}]_{OBJECT.NP}

It is interesting that *which* must be included in (3) but may be omitted from (4). Why is this? In (3), *which* has two functions: (a) introducing the relative clause; and (b) filling the subject slot in the relative clause. We noted that a subject must be stated in each clause. It is in view of property (b) that *which* must be retained in (3); the relative clause must have a subject.

In (4), *which* has the same function (a) but a different (b): referring to the object of the relative clause. As already seen, an object may sometimes be

omitted and this explains why *which* is optional in (4). (The relative clause here does have a stated subject, *the dog*.) Hearing *I saw the cat the dog chased*, a listener knows – from their intuitive knowledge of English grammar – that this is a relative clause construction, with the *which* omitted.

Another facet of grammar is affixes. A regular verb takes one of the inflectional suffixes *-s*, *-ed*, *-ing*, or nothing. There can be derivational affixes; for instance, *pre-Christmas* and *post-Christmas* involve prefixes being added to a noun, *Christmas*, to derive adjectives. These have opposite meanings, ‘before’ and ‘after’, respectively, and a word can include only one of them.

Two prefixes with similar meaning are *super-* (based on *super* ‘over, above’ in Latin) and *over-* (an inheritance from an Old English prefix with the same meaning). Some adjectives take one of the prefixes and some the other: there are *super-natural*, *super-human* and *over-fed*, *over-developed*. A few words can take either prefix, with a slight but significant difference in meaning; *super-* is ‘to a high degree’ and *over-* ‘too much, more than is desirable’. If you move into a new house and describe the neighbours as *super-friendly*, this implies that their behaviour is desirable and useful. But if they are *over-friendly* it implies that they act in a way which you find intrusive and unacceptable. A word may include either of these prefixes, but not both at once.

Grammar works with small systems of terms. Some are affixes and others separate words. There are a limited number of terms in each system, such that each can be identified as not being any of the others.

For instance, English has demonstratives, whose primary meaning is to point to something in the context of speaking: *Would you like this cake?* (pointing to one nearby); *No, I’d rather have that one* (pointing to a cake at the other end of the table). Demonstratives deal with ‘near/far’ and ‘singular/plural’, making up a closed system. ‘I’m thinking of a demonstrative. It is not *this*, *that*, or *those*, what is it?’ The answer is *these*. There are precisely four terms in the demonstrative system; if something is not the first, second, or fourth, it must be the third.

Some of the slots in the grammatical infrastructure of the language are filled by grammatical items, such as demonstratives, articles and pronouns. Others are filled by lexical words – nouns, verbs, and adjectives. Grammatical systems each contain a limited number of forms, and they are closed. That is, new items cannot easily be added to them; no new demonstrative, article, pronoun, interrogative word, preposition, etc.

In contrast, lexical words fall into large open-ended (not closed) classes. New nouns, verbs, and adjectives are constantly being added to the language. A lexical class consists of words with the same grammatical profile. *Apple* and *pear* are both nouns, with the same basic structural possibilities. Similarly for adjectives *clever* and *stupid*, verbs *walk* and *run*, and so on.

Grammatical forms – such as *pre-*, *post-*, *this*, *the*, and *you* – are fully specified in terms of the interlocking structures and systems of the grammar; nothing else remains to be said about them. Complementary to the grammar is a dictionary, which contrasts and explains the meanings of words in the open lexical classes.

Terms in a grammatical system are mutually exclusive; they cannot occur together. That is, a word cannot include the two prefixes *pre-* and *post-*, nor can an NP include the two demonstratives *this* and *that*. There is no such constraint on lexical words. *Large* and *small* have opposite meanings (a bit like *pre-* and *post-*), but it is not too difficult to think up a situation in which they may co-occur. Suppose that Ingrid shows Pablo her collection of small lacquered boxes from the East. Pablo picks up one and remarks: *This is a pretty large small box!*

The dictionary picks things up where grammar leaves off. The grammar is a complex system, with constructions, affixes, and grammatical words. What is the lexicon? Simply a set of words? No, absolutely not. Just like grammar, the lexicon is a complex system, but one of a rather different nature. Each lexical word enters into a series of relationships, linking it to items with similar meanings, with opposite meanings, with compatible meanings but different grammar, and so on.

As a brief illustration, we can take as a starting point the verb *dislike*. Its links include:

- Verbs which are less common than *dislike* and have a more specialised meaning, such as *loathe*, *abhor*, and *detest*. For example, *I loathe insincerity* conveys much stronger emotions than *I dislike insincerity*.
- Verb *hate*, which is more common than *dislike*, and has a wider meaning. Any instance of *dislike* can be substituted by *hate*, producing an acceptable sentence with similar but modified meaning. However, *hate* may not be substituted by *dislike*; not, for instance, in *I'd hate to have to tell your daughter the bad news*.
- Verbs with an opposite meaning to *dislike*, *hate*, *loathe*, *abhor*, and *detest*. They include *like*, *love*, and *enjoy*.
- Branching out from these positive verbs is *prefer*. The sentiment *Mary prefers London to Berlin* implies *Mary likes London more than she likes Berlin*.
- Leading off in another direction from *like*, there is the verb *please* which effectively interchanges subject and object with respect to *like*. One could say either (a) *Robin likes Mary's dancing*, or (b) *Mary's dancing pleases Robin*. A difference is that for (a) Mary may not have been aware that Robin was watching her dance, whereas (b) implies that Mary was doing it on purpose in order to try to please Robin.

This provides a sample of the kinds of linkages which knit together the lexical system.

The next two sections briefly describe two important construction types in English grammar – comparatives and complement clauses. Each will be referred to in later chapters.

Comparative Constructions

The basic comparative construction involves comparing two entities in terms of the degree of a gradable property relating to them. One of the entities is taken as the STANDARD of comparison; the other, which is being compared against the standard, is the COMPAREE. The property is the PARAMETER of comparison and there is an INDEX, showing the type of comparison. For example:

	COMPAREE		INDEX	PARAMETER		STANDARD
(1)	Simon	is	more	famous	than	Gordon
(2)	Gordon	is	less	famous	than	Simon

This is a copula sentence, with copula verb *is*. The comparee is copula subject and the parameter is copula complement, with the standard being marked by preposition *than*.

For negative comparison, the index is always *less*. But for positive comparison the index varies. Some adjectives require *more* before them, as in (1), while others take suffix *-er*. For example, with *quiet*:

	COMPAREE		INDEX	PARAMETER	INDEX		STANDARD
(3)	Max	is	–	quiet	-er	than	Basil
(4)	Basil	is	less	quiet	–	than	Max

A few adjectives do not, by virtue of their meaning, form a comparative; for example *first* and *opposite*. Of the great majority which do, some only occur with *-er*, some only with *more*, while a third set may take either. For example, *Beatrice is clever-er than Doris* and *Beatrice is more clever than Doris* are equally acceptable. This is not a random matter; there are clear principles involved. Whether an adjective takes *-er* or *more* depends on how long its form is, which sound it ends in, and whether it already includes a suffix. The essential principles are:

- I **Adjectives taking *-er*:**
monosyllabics, e.g. *quiet, fat, brave, new, slow, rude*
unanalysable disyllabics ending in vowel /i/, e.g. *heavy, pretty, happy, busy*
- II **Adjectives taking either *-er* or *more*:**
unanalysable disyllabics ending in a vowel other than /i/, e.g. *yellow, clever*
disyllabics which end in suffix *-y* or *-ly*, e.g. *cloud-y, luck-y, friend-ly, man-ly*

III Adjectives taking *more*

disyllabic or longer words ending in a consonant, e.g. *famous*, *superb*,
public, *difficult*, *splendid*

trisyllabic and longer words ending in a vowel, e.g. *ordinary*, *necessary*

No grammar is ever completely tidy and there are a handful of exceptions to these principles – disyllabic adjectives ending in consonants which take *-er*. These include *stupid*, *polite*, and *common*.

There is an explanation for two of the exceptions. If we have a pair of adjectives with opposite meanings, they are likely to behave in the same way.

Clever and *stupid*, *rude*, and *polite* are such pairs. We can see that *clever* and *rude* take *-er* by virtue of their form; it is likely that *stupid* and *polite* also take *-er* by analogy. No similar explanation can be offered for *common*, except that it is a rather common word. This is just an exception. Most things can be explained within a grammar, but not quite everything.

Adjectives form superlatives on the same principles as comparatives, taking *-est* in place of *-er*, *most* rather than *more*, and *least* in place of *less*. (The origins of *more*, *most*, *less*, and *least* are explained in pages 71–2.) There are three adjectives which have irregular forms: *good* with *better* and *best*, *bad* with *worse* and *worst*, and *far* with *farther* and *farthest* (or *further* and *furthest*).

If a dictionary is to fully inform its users how to use words, it should provide information on acceptable comparative indexes for each adjective.

Complement Clause Constructions

For some verbs, subject and object must be noun phrases (NPs); for example, *eat*, *chase*, *wipe*. Others may have either an NP or what is called a ‘complement clause’ in object function. For example:

They will soon announce [the election result]_{NOUN.PHRASE}

They will soon announce [that the Green party has won]_{COMPLEMENT.CLAUSE}

English has three main varieties of complement clause, marked by complementisers *that*, *-ing*, and (*for*) *to*, respectively. They have contrasting meanings.

- A *THAT* clause describes a fact. The complement clause commences with *that*, as in:

I know [that John built the hen-coop]

- An *-ING* clause describes an activity. The verb of the complement clause takes suffix *-ing*, as in:

Fred watched [John build-ing the hen-coop]

- A (FOR) TO clause describes a potentiality (purpose or intention). If complement clause and main clause have different subjects, *for* is placed before the complement clause subject, and *to* before the complement clause verb, as in:

I had intended [(for) Tom to build the hen-coop]

The *for* may optionally be omitted from this sentence. If the two clauses have the same subject then this subject is omitted from the complement clause, together with *for*, as in:

I had intended [to build the hen-coop]

Which complement clause type(s) a given verb occurs with depends on the meaning of the verb and on the meaning of the complement clauses. In chapter 3 there is explanation of why *finish* only takes an -ING clause but *cease* both -ING and TO varieties; and chapter 12 explains why *wish* may take THAT and (FOR) TO complement clauses whereas *want* only occurs with TO clauses.

Some verbs may take all three varieties, with quite different implications. This can be illustrated with *like*.

- *Like* may relate just to the subject's feelings about the fact of a certain thing happening (they may not be at all interested in the internal details of the event). A THAT complement clause is then appropriate, as in:

John likes it [that Mary sings the blues each Friday evening]
 (because she goes out, and he can watch football on TV)

- Or *like* may relate to the subject's feelings about some activity as it unfolds; -ING is then the appropriate complement choice:

John likes [Mary singing the blues]
 (he could listen to her all night)

- Or, the main clause subject might have good (or bad) feelings about the complement clause subject's getting involved in an activity (without necessarily enjoying the activity per se); a (FOR) TO complement will then be used:

John would like [(for) Mary to sing the blues]
 (because he thinks her voice is just right for that style – although in fact his own preference is for opera)

It is interesting to survey the grammatical potentialities for other verbs from the same semantic set as *like*, mentioned a little while back. All can, of course, take a plain NP as object. In addition, there are complement clause possibilities as follows:

- -ING, THAT, and (FOR) TO clauses with *like*, *love*, *prefer*, *hate*
- just -ING and THAT clauses with *dislike*, *loathe*, *abhor*, *enjoy*
- only -ING clause with *detest*

Why, one may ask, should *like*, *love*, *prefer*, and *hate* take a TO clause but not *dislike*? One can say *I'd like/love/prefer/hate [to watch cricket today]*, but not *♦I'd dislike [to watch cricket today]*. The answer lies with derivational prefix *dis-*; it appears that no *dis-* word may occur with a TO complement clause.

Either *Jacob continued [painting the wall]* or *Jacob continued [to paint the wall]* is acceptable, but only *Jacob discontinued [painting the wall]*, not *♦Jacob discontinued [to paint the wall]*. Similarly, *agree* and *allow* can take a TO clause, but *disagree* and *disallow* cannot. It appears that the prefix *dis-* is incompatible with a potentiality sense.

The verb *detest* refers to an activity, not to a fact or a potentiality. For this reason, it can only take as object an NP, or an -ING complement clause; for example, *Mary detests [insincerity]* or *Lucy detests [being photographed]*.

Complement clauses most often occur in object slot, but there are some verbs which accept them in subject function. As mentioned before, *please* reverses subject and object with respect to *like*. Corresponding to the *like* sentences just exemplified, we can have *[That Mary sings the blues] pleases John*, *[Mary singing the blues] pleases John*, and *[For Mary to sing the blues] would please John*.

Plainly, if a dictionary user is to learn how to use a verb appropriately, they must be told what type(s) of complement clauses it may be used with, and in which circumstances.

We now examine the way in which grammatico-lexical information is communicated from speaker to listener – the sound system.

Orthography and Phonology

Mankind has been around for at least 100,000 years (maybe much longer) and for just about all of that time human languages have had sophisticated grammars and extensive lexicons. Writing was invented only about 5,000 years ago. Then, for a long period it was the prerogative of just a few educated people. Until the last few centuries, only a small proportion of people in England habitually wrote more than their name or a brief message (if that).

Today, almost everyone can read and write with fair fluency, but spoken language is still the predominant mode. We talk and listen much more than we write and read. And for most people, the hours spent listening to TV, radio, and films greatly outnumber those spent with a book, magazine, or newspaper.

Speaking is the essence of language use with writing being a secondary manifestation – for communicating when out of earshot, and for producing a permanent record. The purpose of writing is to represent speech through an orthography. In the world today, writing systems differ in how efficient they are. Spanish, for instance, has an almost perfect orthography, with one letter for each contrastive sound, or 'phoneme'. If you can pronounce a word

you know how to spell it, and if you can spell it you know how to say it. For example, the spelling *altura* ‘height’ represents phonological form /altura/. (Slant brackets, /.../, indicate a phonological form, made up of a string of phonemes.) The ways in which phonemes are organised, to make up words and higher units, is called ‘phonology’.

English orthography is of a very different nature. Typically, one phoneme may be written by different letters in different words, and a single letter may represent a multitude of sounds. Consonants pose some difficulty. One letter may represent different phonemes in different words. For example:

- Letter *g* may represent the ‘hard *g*’ sound, /g/, as in *gain* /geɪn/, or the ‘soft *g*’ sound, /dʒ/, as in *giant* /dʒaɪənt/.
- Digraph *th* is used both for the voiced apico-dental fricative /ð/, as in demonstrative *this* /ðɪs/, and for the corresponding voiceless sound /θ/, as in adjective *thin* /θɪn/.

Looking at things the other way round, a phoneme may be represented by different letters in different words. For example:

- Phoneme /s/ is written as *s* in *please* /pli:s/; and as *c* in *niece* /ni:s/.
- Phoneme /z/ is written as *z* in *lazy* /ˈleɪzi/; and as *s* in *rose* /rouz/.
- Phoneme /ʃ/ is written as *sh* in *shoe* /ʃu:/; as *s* in *sure* /ʃuə/; and as *c* in *ocean* /ˈouʃən/.
- Phoneme /ʒ/ is written as *s* in *pleasure* /ˈpleʒə/; and as *g* in *beige* /beɪʒ/.
- Phoneme /tʃ/ is written as *ch* in *chain* /tʃeɪn/; and as *t* in *nature* /ˈneɪtʃə/.
- Phoneme /dʒ/ is written as *j* in *jam* /dʒəm/; as *g* in *gem* /dʒem/; as *d* in *soldier* /ˈsouldʒə/ (or /ˈsouldʒɛ/); and as *dg* in *lodger* /ˈlɒdʒə/.

Vowels are a nightmare in English. The Latin-based orthography used for English has five vowel symbols. But the sound system of Standard British English (what is called ‘Received Pronunciation’ or ‘RP’) has seven short and five long vowels, plus eight vowel sequences or diphthongs. They are, with a sample word for each:

SHORT VOWELS

/ɪ/ as in *bit* /bɪt/
 /e/ as in *bet* /bet/
 /a/ as in *bat* /bat/
 /ɒ/ as in *bottom* /ˈbɒtəm/
 /u/ as in *book* /buk/
 /ʌ/ as in *but* /bʌt/
 /ə/ as in *batter* /ˈbatə/

LONG VOWELS

/i:/ as in *beat* /bi:t/
 /a:/ as in *barter* /ba:tə/
 /ɔ:/ as in *bought* /bɔ:t/
 /u:/ as in *boot* /bu:t/
 /ə:/ as in *Bert* /bɜ:t/