Notes on the Cambridge Grammar of the English Language

Gihan Marasingha

August 28, 2023

1 About this document

These are my notes on *The Cambridge Grammar of the English Language* (CGEL).

2 Basic concepts

2.1 Grammar as a description of English

A grammar of a language is one that aims to 'describe the principles or rules governing the form and meaning of words, phrases, clauses and sentences'. It has two parts **syntax** and **morphology**. The former being concerned with the way in which words form larger units (phrases, clauses, sentences) and the latter with the construction of words themselves.

CGEL also touches on some topics outside syntax and morphology. It examines **semantics**, distinguishing between **grammatical semantics**, this being meaning that follows simply from the language system) and **pragmatics**, this being about meaning in particular situations.

Unlike mathematical definitions, those in CGEL are not stipulative. Rather, they are descriptive. They must be, for the grammar contains no undefined terms.

Thus, where CGEL defines the grammatical category *noun*, it does so by describing a set of characteristics of **prototypical members** (also called the **central** or **core** members) of the category. This is to acknowledge that some words that we class as nouns do not satisfy all the attributes specified in the grammar.

For example, *grape* is a prototypical noun: it denotes a physical object, it has an inflectional form contrast between singular and plural, it functions as the head of a noun phrase. However, the noun *equipment* shows no such inflectional form.

Likewise, in 'race discrimination', the noun race does not act as head of a noun phrase. Instead, it is a modifier of the noun 'discrimination'. However, this does not make 'race' an adjective. The word race does not satisfy many of the other characteristics of adjective. For example, it cannot be used predicatively: 'I am race' is ungrammatical.

Put another way, the grammar is not sufficient, by itself, to determine what is and what is not a noun. This discrimination can only be established by something external: the observation of how we use particular words or phrases in practice.

My thought on the matter is that a grammar is simply descriptive in the same sense that Linnaeus' taxonomy is descriptive. In each case, there is something 'out there' that we are attempting to codify. But the thing itself (be it the relationships between biological species or the construction of English language sentences) has no intrinsic *formal* structure. There will always be 'edge cases' that are difficult to pin down.

3 The components of the grammar

As mentioned, the main components of the grammar are the study of syntax and the study of morphology. We briefly treat each of these in turn.

3.1 Syntax

The theoretical framework for syntax in CGEL is founded on three principles:

1. Sentences have parts, which may themselves be parts.
2. The parts of sentences belong to a limited ranges of types.
3. The parts have specific roles or functions within the larger parts they belong to.

These three principles correspond and lead to the notions of **constituent structure** analysis, **syntactic categories** and **grammatical functions**.

3.1.1 Constituent structure

Consider the foundational notion that a sentence can be divided into parts, each of which may also be divided into parts, called **constituents**.

Figure 1 shows one way to decompose the clause, 'a bird hit the car' as a hierarchy of constituents.

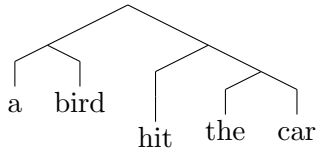

Figure 1: A syntax tree for a simple sentence

In this diagram, the phrases 'a bit' and 'hit the car' are the **immediate constitutions** and the leaf nodes (here words) are the **ultimate constituents** of the sentence.

Evidently, this is not the only possible way to impose a hierarchy on (equally to *analyse*) 'a bird hit the car'. That this is the *correct* analysis can only be appreciated by bringing to bear other aspects of the grammar.

3.1.2 Syntactic categories

We have recognised that sentences can be recursively analysed into 'parts'. We have claimed that each part belongs to a limited ranges of categories. In this section, we briefly examine those categories.

First, we divide the class of categories in two: the **lexical categories** and the **phrasal categories**. Every ultimate constituent (i.e. word) of the sentence belongs to a lexical category. A higher-level structure that depends on a central word and which may contain more than one word is typically a phrasal category.

There are other classes of syntactic structure. These include coordination and supplementation.

Lexical categories The lexical categories consist of the 'parts of speech', these being noun (N), verb (V), adjective (Adj), adverb (Adv), preposition (Prep), determinative (D), subordinator, coordinator, and interjection.

Thus, 'a' is a determiner and 'bird' is a noun. The syntax tree in Figure 2, gives lexical-category labels for each of the words in 'a bird hit the car'.

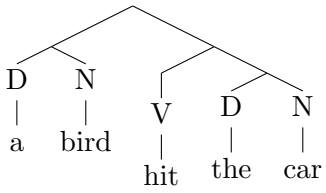

Figure 2: A syntax tree with lexical category labels

Phrasal categories A **phrase** is a constituent at a higher level than the lexical level. Each phrase typically contains a 'most important word' that determines its type. It may also contain other constituents that serve to elaborate its meaning.

For example, a phrase consisting of a noun and supporting constituents is a **nominal**; a nominal together with a determinative is a **noun phrase**; a verb with its supporting constituents is a **verb phrase**; a noun phrase with a verb phrase is a **clause**.

The complete set of phrasal categories is: Clause (Clause), verb phrase (VP), noun phrase (NP), nominal (N), adjective phrase (AdjP), adverb phrase (AdvP), preposition phrase (PP), and determinative phrase (DP).

Figure 3 supplements our previous diagram with the inclusion of phrasal category labels.

4 Grammatical function

The last of our three foundational pillars of syntax is **grammatical function**. The function of a constituent denotes its relationship either to the larger construction containing it to to another element within that construction. The term 'construction' is

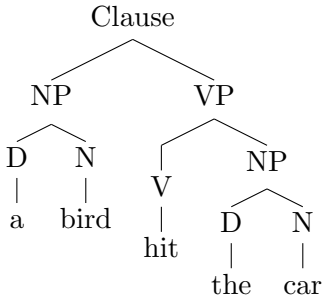

Figure 3: A syntax tree with phrasal and lexical category labels

not defined in CGEL. I take it to mean a phrase when considered as a structure whose elements are related to each other and which may be related to higher-level structures.

The major two types of function are *head* and *dependent*. In the noun phrase 'a bird', the noun 'bird' is the head and the determinative 'a' is a dependent of the construction.

Each construction has at most one (and usually exactly one) **head**, its most important element. Generally, the category of a phrase depends on that of the head: a phrase with a noun as a head is a noun phrase, a phrase with a verb phrase as a head is a clause.

A phrase can contain more than one dependent. Consider the verb phrase 'gave the key to the landlord'. Its head is the verb 'gave'. It has two dependents: 'the key', a noun phrase and 'to the landlord', a preposition phrase.

Depending on the category of the construction at hand, we assign special names to the head and dependents. For instance, a canonical clause has two constituents: a **subject** and a **predicate**. Here, predicate is just a special case of head and subject is a special case of dependent. The subject canonically takes the form of a noun phrase and the predicate a verb phrase.

We see this in Figure 4 where 'a bird' is the subject (dependent) and 'hit the car' is the predicate (head) of the clause. We use a triangle (or roof) over a constituent to show that the analysis is not complete to the lexical level.

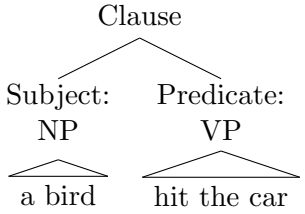

Figure 4: A syntax tree with functional and phrasal labels

A phrase can have more than one dependent. In Figure 5, we have a syntax tree for the verb phrase 'gave the key to the landlord'. The head of a verb phrase is called a predicator and is a verb. Here, the verb 'gave' is the head (predicator). Here the verb phrase has two dependents: one an 'object' and one a 'complement'.

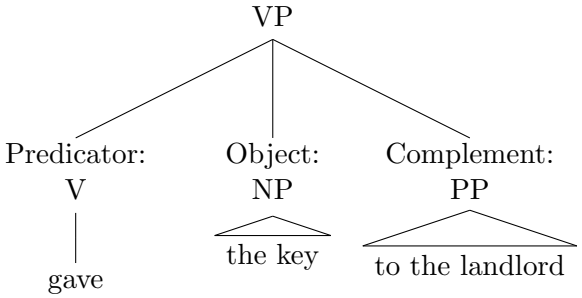

Figure 5: A phrase with two dependents

The previous example shows an important distinction between syntactic categories and grammatical functions. Whereas the phrase 'the key' is always a noun phrase in any context, its grammatical function depends fundamental on the context. In the clause, 'the key is shiny', the phrase 'the key' is a subject. But the same phrase is an object in the clause, 'Bob stole the key'.

Not every construction has a head. The noun phrase 'fish and chips' has no head! In addition to being a noun phrase, 'fish and chips' has the syntactic structure of **coordination**, this being a relation between two or more elements of syntactically equal status, the **coordinates**, usually linked by a **coordinator** such as *and* or *or*.