

Syntax

When we set out to provide an analysis of the syntax of a language, we try to adhere to the “all and only” criterion. This means that our analysis must account for all the grammatically correct phrases and sentences and only those grammatically correct phrases and sentences in whatever language we are analyzing. In other words, if we write rules for the creation of well-formed structures, we have to check that those rules, when applied logically, won’t also lead to ill-formed structures.

For example, we might say informally that, in English, we put a preposition (near) before a noun (London) to form a prepositional phrase (near London). However, if we use this as a rule of the grammar to create structures, we will end up producing phrases like *near tree or *with dog. These don’t seem to be grammatically correct, so we mark them with an asterisk *. We clearly need to be more careful in forming this rule. We might have more success with a rule stating that we put a preposition before a noun phrase (not just a noun).

A noun phrase can consist of a proper noun (London), a pronoun (you) or a combination of an article (a, the) and a noun (tree, dog), so that the revised rule can produce these well-formed structures: near London, with you, near a tree, with the dog. When we have an effective rule such as “a prepositional phrase in English consists of a preposition followed by a noun phrase,” we can imagine an extremely large number of English phrases that could be produced using this rule. In fact, the potential number is unlimited. This reflects another goal of syntactic analysis, which is to have a small and finite (i.e. limited) set of rules that will be capable of producing a large and potentially infinite (i.e. unlimited) number of well-formed structures. This small and finite set of rules is sometimes described as a **generative grammar** because it can be used to “generate” or produce sentence structures and not just describe them.

This type of grammar should also be capable of revealing the basis of two other phenomena: first, how some superficially different sentences are closely related and, second, how some superficially similar sentences are in fact different.

Deep and surface structure

Two superficially different sentences are shown in these examples:

Charlie broke the window.

The window was broken by Charlie.

In traditional grammar, the first is called an active sentence, focusing on what Charlie did, and the second is a passive sentence, focusing on The window and what happened to it. The distinction between them is a difference in their **surface structure**, that is, the different syntactic forms they have as individual English sentences. However, this superficial difference in form disguises the fact that the two sentences are very closely related, even identical, at some less superficial level.

This other “underlying” level, where the basic components (Noun Phrase+Verb+Noun Phrase) shared by the two sentences can be represented, is called their **deep structure**. The deep structure is an abstract level of structural organization in which all the elements determining structural interpretation are represented. That same deep structure can be the source of many other surface structures such as It was Charlie who broke the window and Was the window broken by Charlie?. In short, the grammar must be capable of showing how a single underlying abstract representation can become different surface structures.