Phrase Structure Grammars

Scott Martin

November 6, 2008

In addition to representing the syntactic structure of a language in a simplified, condensed way, phrase structure grammars importantly also capture two of the more interesting aspects of natural language syntax: ambiguity and recursion.

Ambiguity

The English sentence in example (1) is ambiguous:

(1) Mary saw the dragon in the cave.

In this case, ambiguity arises from the fact that there are (at least) two ways to interpret the sentence's meaning. In one reading, the speaker is in the cave (call it **reading A**), while under the second reading the dragon is in the cave but the speaker is not (**reading B**).

Phrase structure grammars can capture ambiguity by assigning more than one structure to a given string. Below are the phrase structure rules that would license the sentence in (1):

S → NP VP	
NP → Det N	Det → *the*
NP → NP PP	N → *cave*
NP → Pro	N → *dragon*
VP → V NP PP	P → *in*
VP → V NP	NP → *Mary*
PP → P NP	V → *saw*

Table 1: Phrase structure rules licensing the sentence in (1).

Note that the grammar in Table 1 contains **two** rules with the category VP on the left. This means that there are two structures that this grammar can categorize as being VP-type things. This is where the ambiguity will be captured in the case of item 1.

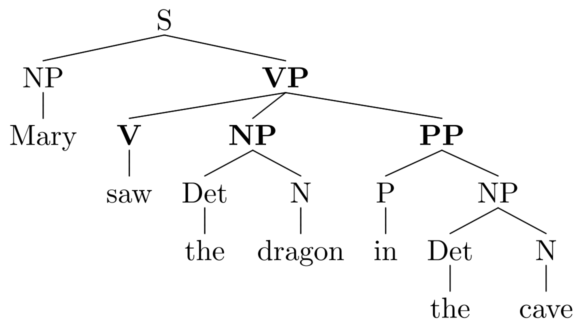

Figure 1: Phrase structure tree for reading A of (1).

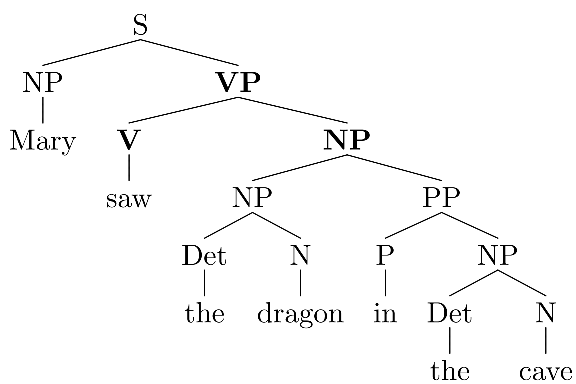

Figure 2: Phrase structure tree for reading B of (1).

To see this in action, consider the phrase structure trees in Figures 1 and 2. In Figure 1, the phrase structure rule V $\rightarrow$ V NP PP is used. In this case, the PP constituent is a top-level sub-constituent of the VP because it is describing where the V constituent *saw* happened (in the cave). However, the tree in Figure 2 uses the V $\rightarrow$ V NP rule, where the NP is the direct object to the verb. Here, the NP contains the PP, whose role is to describe something about the NP (that it is in the cave). This crucial differences between Figure 1 and Figure 2 is highlighted using bold text. The phrase structure rules here correctly describe the two ambiguous readings by assigning different structures for the VP. It is important to note that the grammar described in Table 1 is capable of capturing both readings of the ambiguous sentence in item 1.

Recursion

Imagine that we extend the grammar in Table 1 with the rules in Table 2. These simple additions, essentially composed of two new prepositions and

P → over
P → with
N → hill
N → telescope

Table 2: Additional phrase structure rules extending the grammar in Table 1.

two new nouns, allow the grammar to capture the recursive embedding illustrated in the following examples:

(2) a. Mary saw the dragon in the cave.
b. Mary saw the dragon in the cave over the hill.
c. Mary saw the dragon in the cave over the hill with the telescope.

(Example (2a) is repeated from (1).) To see how the grammar handles this recursive building, consider the two rules that describe prepositional phrases: NP → NP PP and PP → P NP. Looking at either rule we see that an element of the right side of the rule is the left side of the other rule. For example, the rule PP → P NP contains an NP in its right side, while NP is the left side of the rule NP → NP PP. Likewise, NP → NP PP contains a PP in its right side, and this is the left side of PP → P NP. Since each rule contains the left side of the other in its right side, infinitely large structures can be built from the sequences they describe.

The following examples demonstrate this recursion, with the recursive rules highlighted. To save space, only the relevant parts are shown.

Figure 3 shows a structure that uses both of the recursive rules in the grammar in Table 1 in a single embedding. In this most basic case, the possible repetition of structures via recursive rules does not occur.

In Figure 4, the NP *the cave* is replaced with an expanded NP *the cave over the hill*, which is constructed using the rule NP → NP PP. Here, the

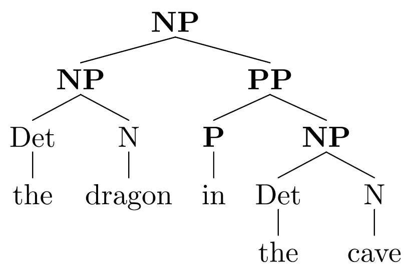

Figure 3: Phrase structure tree for (2a), repeated from Figure 2.

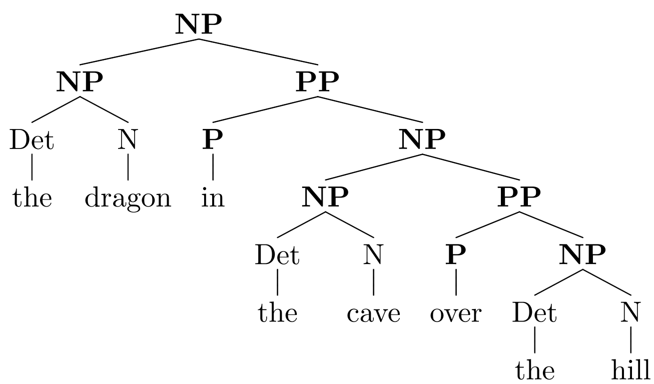

Figure 4: Phrase structure tree for (2b).

recursion occurs once since both recursive rules NP → NP PP and PP → P NP are used exactly twice.

The sentence in Figure 5 shows a continuation of the recursion pattern. Starting with the sentence in (2b), the NP *the hill* is expanded to the NP *the hill with the telescope* by invoking the same rule as was used to expand the NP in Figure 5.

In each of these examples of the recursion described by our phrase structure grammar, an NP built using the simple rule NP → Det N is expanded to an NP containing a prepositional phrase by invoking the rule NP → NP PP. It is important to note why this is key: since this rule contains a PP on its right and the rule PP → P NP in turn contains an NP on its right, the NP expansion shown in these examples can recur infinitely many times.

Recursive structure is present in many places in natural language syntax. One of the others is the embedding of sentential complements, as in (3)

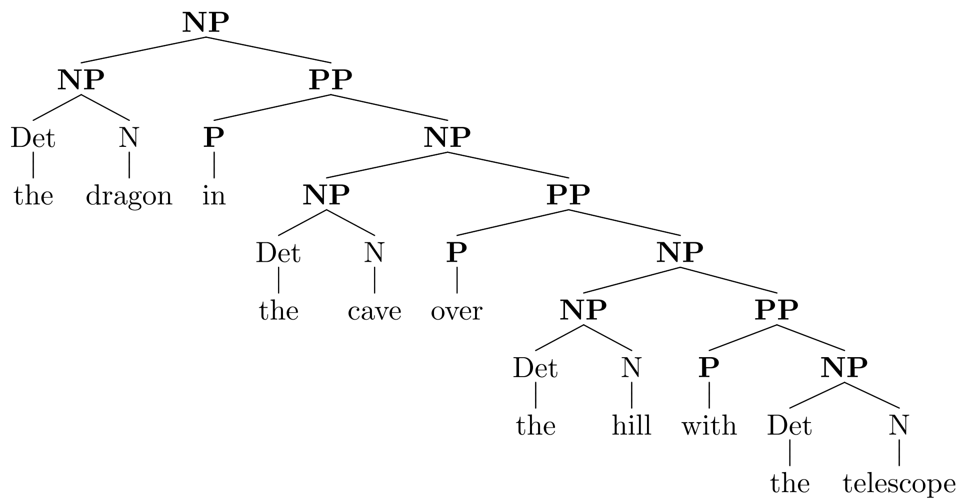

Figure 5: Phrase structure tree for (2c). This analysis reflects a reading of (2c) where the telescope is on the hill.

(3) a. Mary saw the dragon.
b. John knew Mary saw the dragon.
c. Susan thought John knew Mary saw the dragon.
d. Bill believed Susan thought John knew Mary saw the dragon.

This recursion stems from the fact that verbs such as *know*, *think*, and *believe* take sentences as their complements. Adding a subject then creates a new sentence, and the process can repeat again with a higher-level sentential complement verb.

Another example is that of coordinating conjunctions like *and*. These conjunctions take two like things (say, two NPs) and create a new thing of the same type. For example, a grammar that deals with coordination might contain rules like the following: Such rules would allow an analysis

NP → NP Coord NP
Coord → *and*

Table 3: Phrase structure rules for coordinating NPs.

of NPs that grow recursively using coordination, like the examples in (4) (these examples assume all proper names are NPs):

(4) a. Mary slept.
b. Mary and John slept.
c. Mary and John and David slept.
d. Mary and John and Paul and Susan slept.

All [1] of the subject NPs in (4) share a common structure where each NP immediately contains two other NPs. Figure 6 demonstrates this application of the rules in Table 3. The vertical ellipses in Figure 6 show nodes where

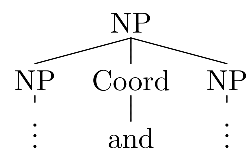

Figure 6: General structure of coordinated noun phrases, as licensed by the rules inTable 3.

the structure it depicts could be repeated.

[1]Besides the top-level NP, which in the case of (4a) is the *only* NP.