TUTORIAL
5.12

Syntax 1

This tutorial and the next one introduce syntax - the unconscious principles people use to put words together into sentences. We will explain and describe what syntax is, and how to recognise and represent the ways that sentences are built, using examples from English and other languages.

Review English Grammar

It would be a great idea for you at this point - if you need to - to take some time to do a review of basic English grammar. Your goal would be to become familiar with the basic structure of English and the terminology used to describe it. This will give you a good foundation for the next tutorials which introduce the variety of possible structures that you may encounter in other languages. There are many resources online for reviewing English grammar which are easy to use and that you might even find enjoyable.

What is syntax?

Knowing a language doesn't just mean knowing all its sounds and words. It also means unconsciously knowing the principles for how words are put together to build sentences. These sets of unconscious principles are called the *syntax* of the language. If a speaker uses the correct rules for connecting words together, it is possible for them to create an infinite number of sentences, all of which are meaningful to a listener who also knows the same rules.

The rules of grammar do not have to be *explicitly* understood by the speaker of the language or the listener. The majority of native speakers of a language will have no formal knowledge of the grammar of language but are still capable of speaking the language grammatically to a great degree of accuracy. Native speakers of a language assimilate these rules subconsciously while the language is being learned as a child.

As we said before, the grammar of a language has several components:
– The *phonetics* that governs the structure of sounds;

- The *morphology* that governs the structure of words;
- The *syntax*, which governs the structure of sentences;
- The *semantics* that governs the meanings of words and sentences.

Since we have already investigated phonetics and morphology, now we are going to focus on syntax and later we will begin to look into semantics.

Phrase structure and argument structure

There are two main areas of syntax that overlap and are connected: phrase structure and argument structure.

The *phrase structure* of a language is the set of principles about how words are put together to make phrases, and how words and phrases are put together to build sentences. Phrase structure is also often called constituent structure, because it deals with the way the constituents (or separate parts) of a sentence are put together.

Argument structure is about the way the participants in an event are expressed in the grammar of a sentence (here, the term 'argument' doesn't mean a disagreement). In the sentence *Mary saw the dog*, the event has two participants (or arguments): Mary and the dog. One is the subject of the sentence, the other is the object. One is the participant who does the seeing, the other is the one who gets seen. The way a language tells us which is the subject and which is the object and so on, is called argument structure.

The syntax of a language is made up of the rules for phrase structure and the rules for argument structure.

An example of the phrase structure in English is the fact that most adjectives come before the noun they are telling us more about, e.g.: *the red house, the big dog, the yellow sun.* But other languages have a different phrase structure, for example in French most adjectives come after the noun: *la maison rouge* (the house red). This is also true for Ata (PNG): *tani lauoinu* (the house red).

Possible sentence structures

The main parts of a sentence are the subject (S), object (O) and verb (V). Different languages use different orders for these three parts in sentences. There are six possible orders: SVO, SOV, VSO, VOS, OSV and OVS, but most languages - over 80% - have the subject of the sentence first. Only a very small number have the object of the sentence first.

Also, over 90% of languages have the verb and object together. This makes sense, because the relationship between a verb and its object is particularly important. Why? Because if the verb and object are next to each other there

can be a Verb Phrase, and every sentence must have a Noun Phrase and a Verb Phrase to make sense.

We are first going to look at some examples from English to show you how to diagram and discover how syntax works in your own language, then later we will take a look at some examples in other languages.

Phrases

Sentences are made up of smaller phrases. There are several different types of phrase that can be used in a sentence, but the two phrases which *must* be used in a sentence for it to make sense are a **noun phrase** and a **verb phrase**.

In a phrase, we *must* have a word which is called the **head**. This is the core of the phrase, what the phrase can't exist without. So in a phrase like 'the dog' or 'ran far away', in the first phrase 'dog' is the head because it is the main part of the phrase, and in the second phrase 'ran' is the head because it is essential for the phrase to exist. We can have 'dog ran', which isn't grammatical, but this still makes sense since we can understand that the dog ran. But we can't have 'the far away', this makes no sense to us!

Modifiers

So 'the' and 'far away' have to be given a name to distinguish them from the head. We call these **modifiers**. They modify the head and give it specific meaning. The **determiner** 'the' modifies the 'dog' because it lets us know which dog we are referring to. The phrase 'far away' modifies the verb 'ran' by letting us know the extent to which the dog ran.
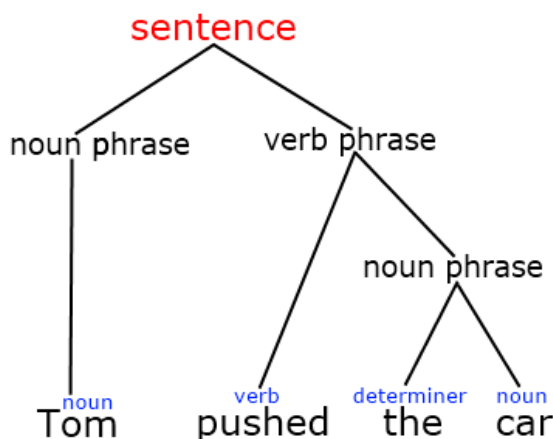
Noun Phrases

A noun phrase is usually the person or thing that is performing the verb in the sentence. It may also be the person or thing that the verb is being done to in a sentence.

The person doing the verb in a sentence is known as the **subject**. For example, in the sentence 'Tom pushed the car', 'Tom' is the subject of the sentence as he is pushing the car. 'The car' is the object in the sentence as the car is the object that the verb is being done to. Both of these are noun phrases.

A noun phrase has to contain a **noun**, such as a name or a tangible object. Sometimes, a determiner is also needed in a noun phrase, for example '*a* cat', '*the* dog' - *'a'* and *'the'* are called determiners because they tell us which person or thing is involved in the sentence.

3

Syntax Trees

To explain sentence structures, we draw trees, like the one above. It is actually a lot more complicated than it looks, so we will explain it step by step. We start with the noun phrase, *the car*. We know a noun phrase is made up of a determiner and a noun. First, we label what parts of speech each of these are:

Next, we can draw two lines to join them together, creating a noun phrase:
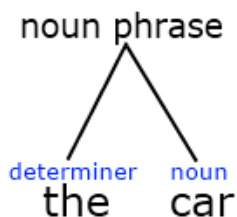
If we look at the proper noun, *Tom*, we find it is a little bit different. A proper noun does not need a determiner, so we can go straight to making it another noun phrase:
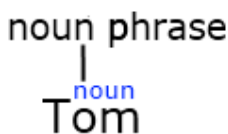
Now, we have drawn noun phrases for both 'Tom' and 'the car' in our sentence. But we aren't quite finished yet, now we need to talk about *verb phrases*.

Verb Phrases

Now that we have made some noun phrases, we can move on to verb phrases. The good news is a basic verb phrase can be made up of one word. The bad news is not every sentence has a basic verb phrase. For now let's look at the basic verb phrase and how the verb phrase fits in. First, we identify the verb and label its part of speech.
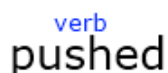
Next, we can label it as a verb phrase. This is connected to the noun phrase 'the car', so we connect them and write the label above. But why is this the case?
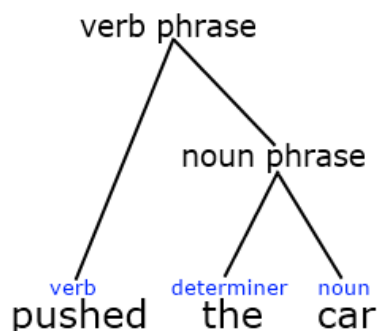
We need to link all of this together. In this sentence, we must join up all of the phrases to make a tree that says 'Sentence' at the top. But we can't just simply join up the lines, we need a structure! In order to figure out the correct structure, we need to look at constituents and how they are related to one another.

Constituents: words and phrases

In most languages, words can't just be put together in any order. For example, look at this sentence: *Last Tuesday the big fat cat chased the small ugly dog.*

You can understand the meaning of that sentence because of the way the words are grouped together into phrases - in the way that English grammar requires. But if it is written like this, it no longer makes any sense: *Small big chased Tuesday dog ugly fat the last cat the*

This second sentence is ungrammatical. For it to be grammatical, the words
the, *big*, *fat* and *cat* have to group together in a particular way to form a
particular kind of phrase.

Words are the most basic building blocks that are used to build sentences - they are the component parts, or *constituents*, of a sentence, and words can also group together into larger units called phrases. So, groups of words (phrases) are also *constituents*. Syntax uses *constituents* as building blocks, and constituents can either be words or phrases.

Syntax does not have a flat structure where words just follow each other in a row. Words join with other words to form phrases, and they in turn join with other words and phrases to form bigger phrases, and so on. So syntax has a hierarchical, nested structure, where constituents combine with other

constituents to form bigger constituents, which combine with other constituents to form even bigger constituents, right up to sentences. That is why it is helpful to draw syntax trees, because they represent this hierarchical structure.

In our example above, *big* and *fat* combine to form the phrase *big fat*, which then combines with *cat* to form the phrase *big fat cat*, which joins with *the* to form the phrase *the big fat cat*, which then combines with *chased* to form the phrase *the big fat cat chased*, etc.

To put it the other way around, constituents like sentences contain smaller constituents such as phrases, and these in turn contain even smaller constituents such as words and phrases, all the way down to each individual word.

Tests for constituency

So, what is a constituent? A constituent is a word or a group of words that function together as a unit. We can figure out fairly easily what words or groups of words are constituents by performing tests on them. There are four tests which can be used on any suspected constituent. These are:
1) *Omission* – Create the same sentence but take away the word/words we are testing.
2) *Replacement* – Can we replace the word/words with just a single word?
3) *Standalone* – Can we form a question using the sentence, and the answer be the word/words being tested?
4) *Movement* – Can we move the position of the word/words in the sentence and have the sentence still make sense without changing the meaning?

Let's do some examples of these tests so you can see how they work.

First, *omission*. If we take the sentence 'He sat down in the car', we can use the omission test to figure out the constituents in the sentence. Let's take 'in the car' as our constituent. If we take this away, does the sentence still make sense grammatically?
> 'He sat down in the car.'
> 'He sat down.'

The sentence still makes sense. Therefore, 'in the car' is likely to be a constituent; however we can't be certain yet. It is always important to do more than one constituency test to figure out if the word/words are constituents. It could pass one and fail the other three!

Next, let's do *replacement*. We take 'He sat down in the car' and replace 'in the car' with just one word. Useful words to try and use in a replacement test are *pronouns* (he, she, it, this, that etc).

'He sat down in the car.'
'He sat down there.'

The sentence makes sense once again! So 'in the car' has passed two constituency tests, but let's try another test first, just to make sure.

Time for the *standalone* test. We need to form a question using the rest of the sentence, and have the answer only be the word/words we are testing.
'Where did he sit down?'
'In the car.'

Success! It surely must be a constituent now. But we'll do the final test just in case. The final test is *movement*. Can we move the words we are testing to see if the sentence still makes sense?
'He sat down in the car.'
'It was in the car that he sat down.'

Once again the sentence makes sense. Now we have done these constituency tests, we know that 'in the car' is a constituent of the sentence 'He sat down in the car.' But be careful, *just because it is a constituent here does not mean it is a constituent in every other sentence*.

Relationships in Trees

So now we can identify constituents, let's go back to our tree. There are several important relationships in a syntax tree that we need to know about.

Domination – phrases can dominate words in a sentence tree. But how do we know which phrases dominate what? We said that a syntax tree has a **hierarchical** structure, which means that one label dominates all that is below it. Look at the picture below:
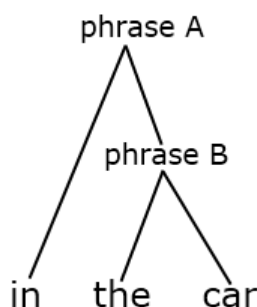
We can see here that phrase A dominates phrase B and 'in'. We know this because phrase A is above phrase B and it is above the word 'in'. When a phrase is directly above something, we call this **immediate domination**. Phrase A immediately dominates phrase B and 'in', but it does not immediately dominate 'the' or 'car'. But it is important to remember that even though it

doesn't *immediately* dominate 'the' or 'car', it still dominates 'the' and 'car' because it is above them in the tree structure.

Naming Phrases

So, we looked before at verb phrases and noun phrases, but these aren't the only types of phrases you can have in a tree. We can also have adjective phrases, adverb phrases and prepositional phrases. We'll quickly talk about how these are different to the other phrases.
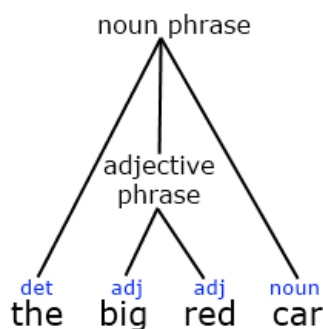
Adjective Phrases

Since phrases are named after the head, adjective phrases are simply phrases where the head of the phrase is an adjective. So for example the sentence 'The big red car drove away' contains the phrase 'The big red car'. This includes the adjectives 'big' and 'red'.

An adjective phrase can contain more than one adjective, so we can make 'big red' into a phrase.
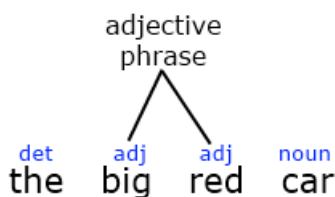
Next, we join this adjective phrase to the noun 'car' to make a noun phrase. But that's not all, since we have to add the determiner 'The' to complete the phrase. From there we can continue to build up the tree.

Adverb Phrases

These work a lot like adjective phrases. They are found near to verbs (because adverbs describe the verb), such as 'immediately' in the sentence 'he sat down immediately'. They are very easy to recognise, and often form adverb phrases containing just one word: the adverb itself. But be careful, adverbs can appear in a sentence and *not* be an adverb phrase! They are often used to modify an

adjective, such as 'the immediately recognisable man'. Here, the adverb 'immediately' does not form an adverb phrase, but modifies the adjective in the adjective phrase 'immediately recognisable'.
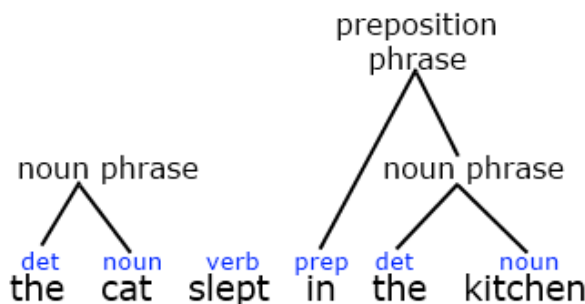
Prepositional Phrases

Another type of phrase that is very common in sentence structures is the prepositional phrase. They may seem complicated at first but they will get easier to recognise with some practise. If we take the sentence 'The cat slept in the kitchen', we can break it down and analyse it to see where the prepositional phrase is.

Having labelled each word with its part of speech, we can see we have the preposition 'in' in this sentence. So let's start putting it together. Hopefully by now, you can see that 'the cat' is a noun phrase, and you can connect it. You should have seen that 'the kitchen' is also a noun phrase, and you can join it up as well.
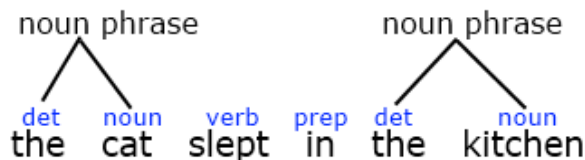
Okay, so now let's join up 'in'. When we connect 'in' and the noun phrase 'the kitchen', we create a new phrase. This phrase is immediately dominating the noun phrase and the preposition 'in'. Therefore, we have made a preposition phrase. The noun phrase 'the kitchen' is modifying the location of where the cat slept, meaning it is modifying the word 'in'. So we know that 'in' is the head of this phrase, making it a preposition phrase.

But we aren't quite finished yet! The preposition phrase 'in the kitchen' modifies the verb 'slept'. So we now join up the verb 'slept' and the preposition phrase 'in the kitchen' to make a new a phrase. Can you guess what phrase we make here?
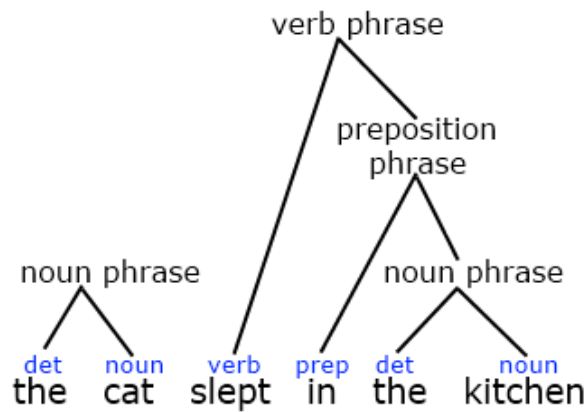
Yes, we make a verb phrase. This is because the verb is the head of the phrase and 'in the kitchen' modifies it. So now we have a noun phrase and a verb phrase left, and we can join them together to make a sentence.
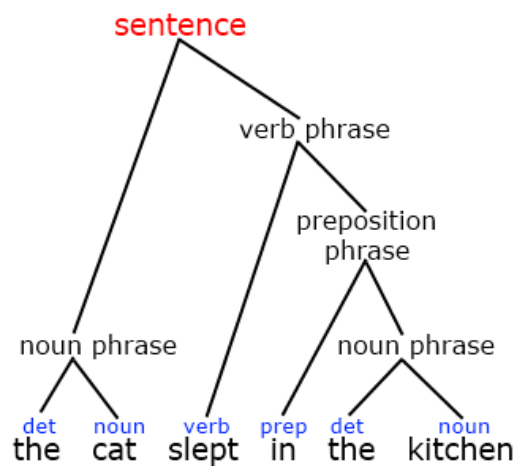

Usually when you want to join a word to a phrase, the word you are adding will be the head of the phrase, meaning that whatever part of speech it is, the phrase will be named after that part of speech. So if we join a noun to an adjective phrase, we will make a noun phrase. And if we join a verb to a noun phrase, we will make a verb phrase.

Co-ordinate Phrases

Co-ordinate phrases are rare phrases that do appear in English. You will have used them already, when you say things like 'fish and chips', 'Max and Ben', and pretty much any situation where you join two nouns together using the **conjunction** 'and'. So how do we deal with these?
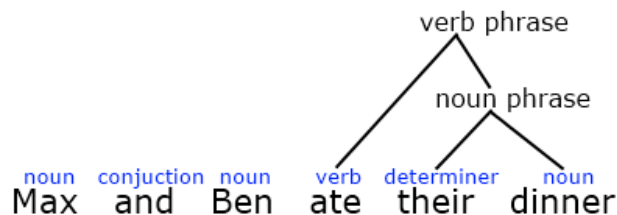
Here we have the sentence 'Max and Ben ate their dinner.' We have labelled the words, and we can construct the tree for the verb phrases, but we need to make a noun phrase for 'Max and Ben'. So how do we make the noun phrase? Well, 'Max' and 'Ben' are both proper nouns, meaning that they don't need a determiner. So here we just simply join 'Max', 'and', and 'Ben' together at the same time.
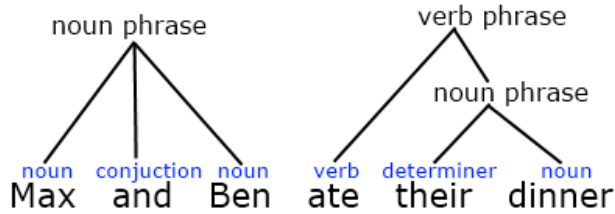

We can then label this as a noun phrase and join it to the verb phrase as a sentence.
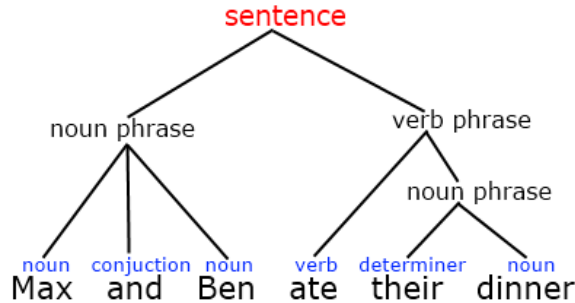
But what if we don't have proper nouns? Let's take the sentence 'The boy and the girl ate their tea.' Again, we form the verb phrase, but what do we do to make the noun phrase? If you can work it out now, write it down and then check your answer with the tree below.
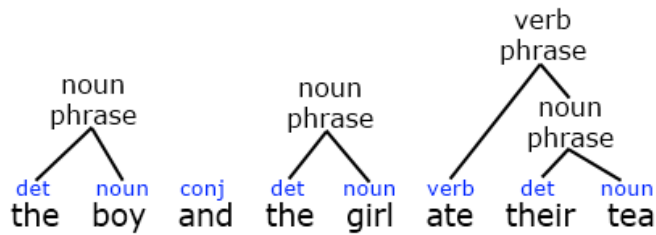

We created two separate noun phrases. We have the noun phrase 'the boy' and the noun phrase 'the girl', with 'and' sitting in the middle of them. Now, we can join them up like we did with the proper nouns, making a noun phrase that we can use to complete the sentence!
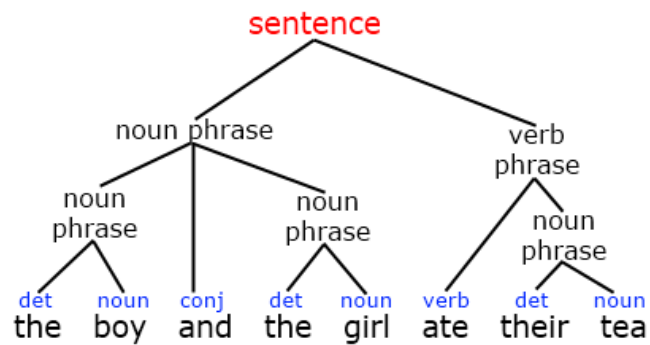
ACTIVITIES
Syntax 1

1. For the following phrases, name the head of the phrase:
    a. in the grass
    b. the cranky old woman
    c. ran as fast as he could
    d. jumped the fence
    e. that delicious meal
2. Using what you have learned, try to draw syntax tree diagrams for the following sentences.
    a. The dog ate the bone.
    b. My brother won the lottery.
    c. The boy kicked the ball.
    d. The dog saw a man in the park. (this one contains a Prepositional Phrase)
    e. The boy lives in the old house by the road. (this one is a bit more difficult, have a try at it, working systematically, but don't worry if you have to peek at the answer!)
3. Invent your own English sentence for the following syntax tree (answers will vary):
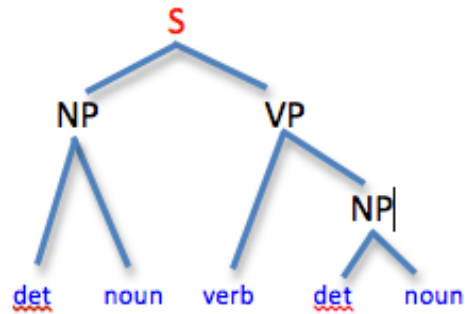
4. Determine if the following languages are VSO, SOV, OSV, etc.

Turkish:
Yusuf   elmayı   yedi
Yusuf the apple ate (Yusuf ate the apple.)

Arabic:
Qara'a l-mudarrisu   l-kitāba.
Read   the teacher   the book (The teacher read the book.)

Hungarian:
Pista kenyeret szel
Pista bread   slices (Pista slices bread.)

Biblical Hebrew:
Vayidaber YHWH el-Moshe
spoke   YHWH to Moses (YHWH spoke to Moses)

Telegu:
Rāmuḍu baḍiki veḷtāḍu
Ramu   to school goes (Ramu goes to school.)

Apurinã:
anana   nota apa
pineapple   I   fetch (I fetch a pineapple.)
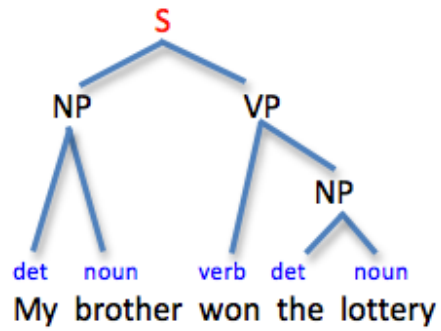
Answers:

1. The head of the phrase:
   a. in the <u>grass,</u>
   b. the cranky old <u>woman,</u>
   c. <u>ran</u> as fast as he could,
   d. <u>jumped</u> the fence,
   e. that delicious <u>meal</u>
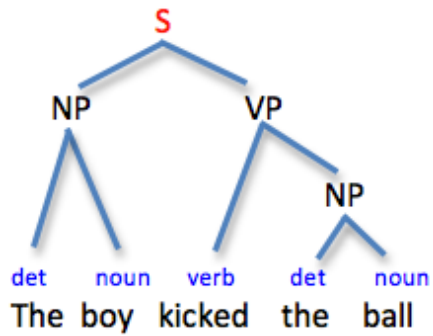
2. Syntax tree diagrams:
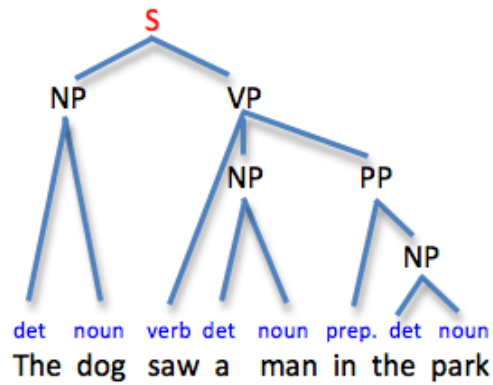   a.   The dog ate the bone.
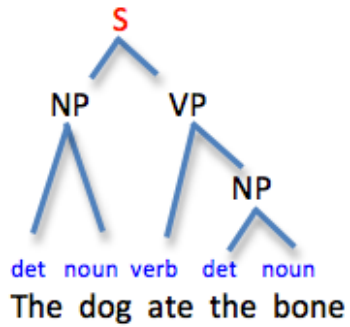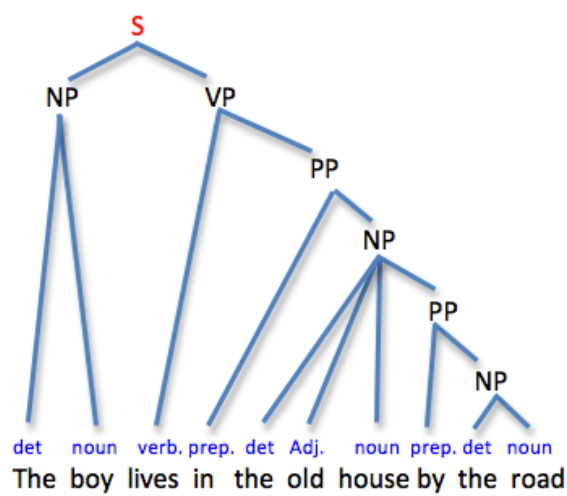

    b.   My brother won the lottery.

    c.   The boy kicked the ball.

    d.   The dog saw a man in the park.

e. The boy lives in the old house by the road.

3. Answers will vary.

4. Sentence structure:

   Turkish: SOV. **Arabic**: VSO. **Hungarian**: SOV.
   Biblical Hebrew: VSO. **Telegu**: SOV. **Apurinã:** OSV.