

Syntax

John Goldsmith

August 26, 2013

Syntax

It has long been recognized by linguists that the construction of a sentence is more than stringing a set of words together: there is a structure to it, one which is not usually indicated in the written form of the language but which is there for us to analyze. ¹ Starting in the 1940s, American linguists used ambiguous sentences — strings of words with two obviously different analyses—to drive this point home. Here are some examples of that; headlines are particularly good sources of funny ambiguous sentences: ²

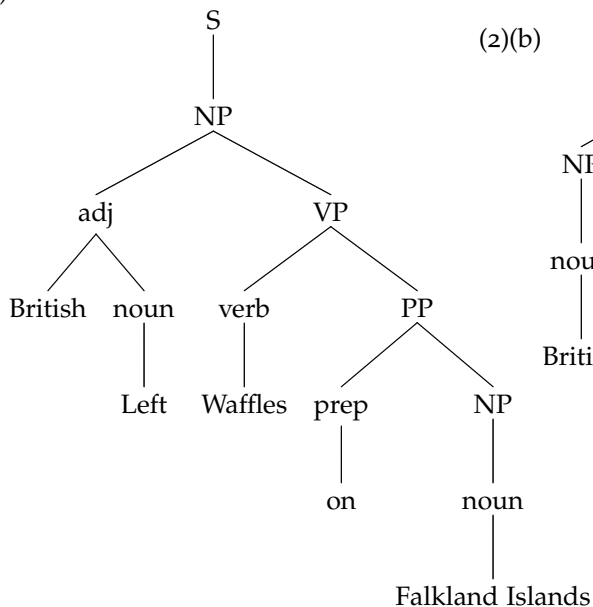
¹ Thanks to Jason Merchant for comments on an earlier version.

² thanks to the morphology book by Mark Aronoff and Kirsten Fudeman.

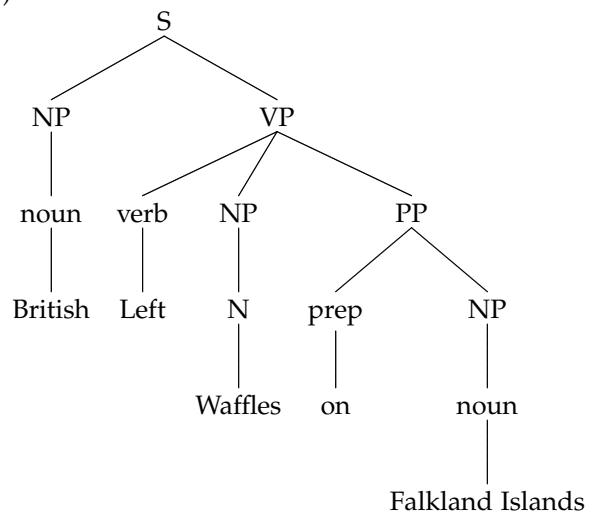
British Left Waffles on Falkland Islands.
Miners Refuse to Work after Death.
Eye Drops Off Shelf.
Local High School Dropouts Cut In Half.
Reagan Wins on Budget, But More Lies Ahead.
Squad Helps Dog Bite Victim.
Juvenile Court to Try Shooting Defendant.
Kids Make Nutrious Snacks.

We will develop a method that will generate two analyses for these sentences, like the two below for the first example above:

(1)(a)



(2)(b)



Phrase structure rules (PSR)

the goal of syntax is to understand how we put words together to create well-formed, and meaningful, sentences. It is clear right from the start that we are looking at sequences of words: words occur one after another, in sequence. What are the principles governing the relative order of words in sentences? Until the middle of the 20th century, thinking about this problem divided into two methods: in the first, individual words would be identified in the sentence by the role they played in a sentence. For example, in the sentence: *Lee sent a birthday present to Kim*, *Lee* is the subject, *present* is the direct object, and *sent* is the verb. In the second approach, the sentence would be broken up into smaller and smaller pieces.

In the mid 1950s, this second analytic approach was stood on its head, and linguists began to write synthetic rules that generated pieces of sentences. These pieces could be as simple as a word, or it could be very complex. These rules were formulated—first by Noam Chomsky—in a way that was inspired by mathematical logic. For example,

(3) $S \rightarrow NP VP$

is a rule that says that an S[entence] can be expanded as an NP (a Noun Phrase) followed by a Verb Phrase. And we will have to immediately write some other rules to provide an answer to what those things are. We will expand VP in this way:

(4) $VP \rightarrow verb NP$

and we will expand NP in this way:

(5) $NP \rightarrow det adj noun$

We will distinguish between *lexical categories*, such as noun, adj[ective], and det[erminer], and *phrasal categories*, such as S, NP, or VP (sentence, noun phrase, and verb phrase). Lexical categories are the most specific things that our syntax will delve into, at least at the beginning; and our phrase structure rules

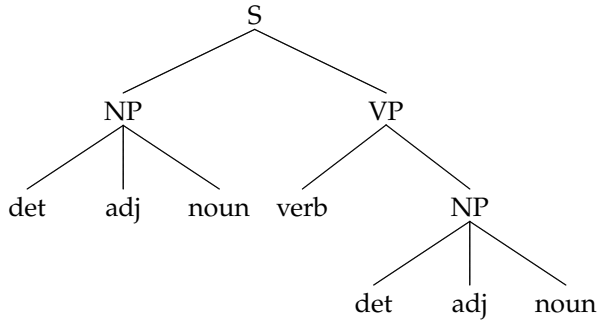
We begin with an initial symbol (for now, S), which is expanded by means of phrase-structure rules, until the bottom categories of the tree that is created consists entirely of lexical categories; these lexical categories then are filled out with lexical items of the appropriate category (nouns, adjectives, and so on).

We will use lower case letters to specify lexical categories: this is not standard notation, but it is convenient.

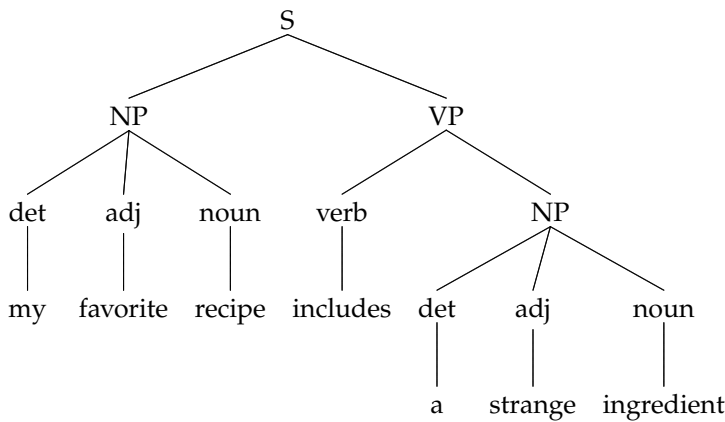
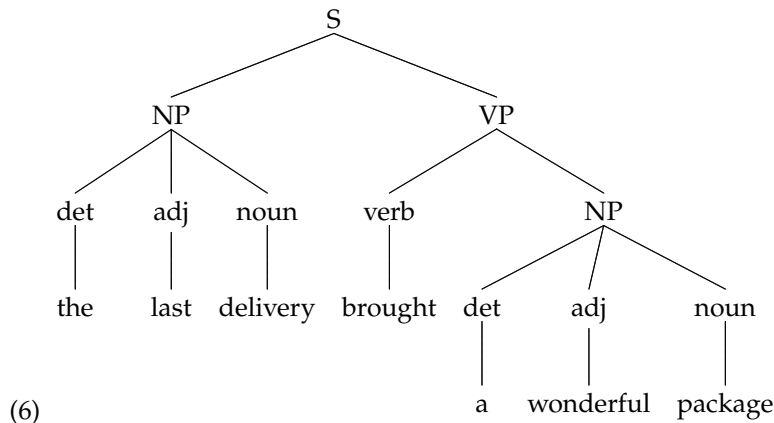
We could write successive expansions in this way:

expansion	the operative rule
S	
NP VP	$S \rightarrow NP VP$
det adj noun VP	$NP \rightarrow det adj noun$
det adj noun verb NP	$VP \rightarrow verb NP$
det adj noun verb det adj noun	$NP \rightarrow det adj noun$

but it is much more common to draw this as a tree:



And this tree represents many millions of sentences, two of which are drawn here:



Big Idea: the motivation for positing the rule $NP \rightarrow \text{det adj noun}$ is that this sequence appears several times in the description of the English sentence, and we can make the overall description more compact if we posit this entity, the 'NP'.

The more times we are able to simplify our overall description by re-using a phrasal (non-lexical) category like *NP*, the better we believe our analysis is motivated. So, for example, there is another VP-expansion that is motivated by examples like *send a big present to the new teacher*. Instead of accounting for this with a new VP-expansion rules

(7) $VP \rightarrow NP \text{ prep det adj noun}$,

we write instead:

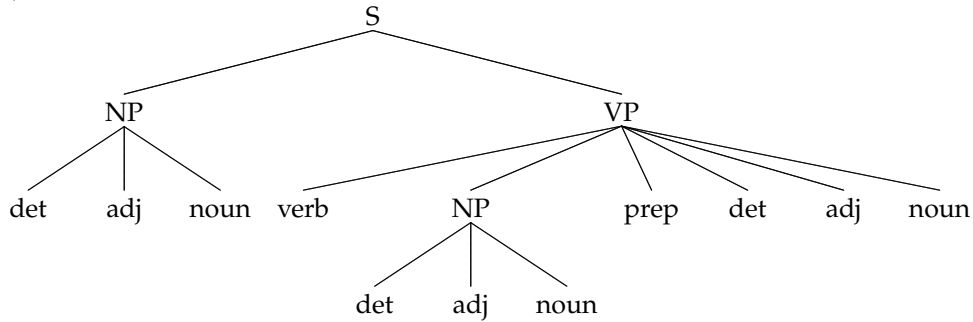
(8) $VP \rightarrow NP PP$

(9) $PP \rightarrow prep NP$,

where *prep* is a lexical category of prepositions that includes such words as *to*, *for* and *with*, and 'PP' marks a prepositional phrase.

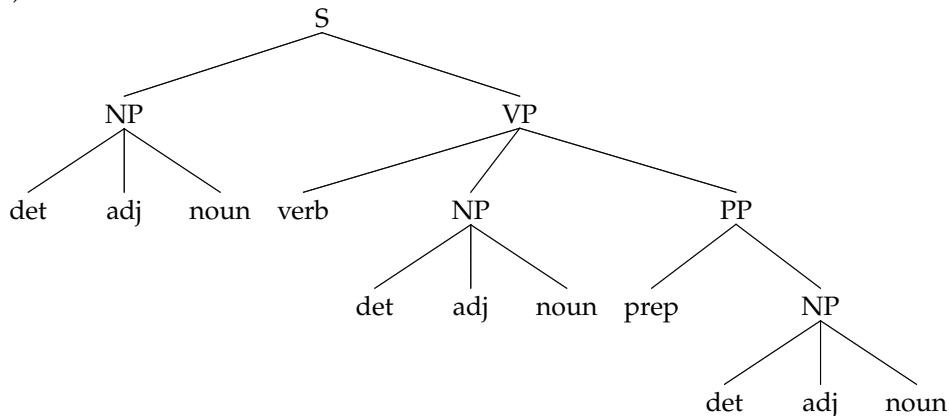
Thus the tree structure is not:

(10)



but rather:

(11)



Alternative expansions of phrasal categories

We have just noted that there are two possible expansions for VP:

(i) verb + NP and (ii) verb + NP + PP. In general, phrasal categories do have a lot of different, but related, ways of being expanded, and this fact is a central part of the motivation for talking about phrasal categories in the first place. Let us explore this.

Now, there is an implicit *independence assumption* made when we posit a category such as NP or VP: no matter where that node is generated by phrase-structure rules, any of its expansions may appear in that position. There is a lot that is right about that assumption; but it is by no means the whole story, and to be perfectly blunt about it, it is far from true: it is, indeed, false. False but helpful.

For example, let us consider several possible expansions for NP in English:

Perhaps the first reference to this is in Pittman 1948: if we do not view a sentence as being hierarchically broken into parts, "one is almost compelled to regard every morpheme in an utterance as pertinent to the description of every other morpheme. But a good analysis in terms of immediate constituents usually reduces the total possible environmental factors of a given morpheme or sequence of morphemes to one: in other words, it states that the only pertinent environment of a given immediate constituent is its concomitant (the other immediate constituent)." (p. 287)

(12)

- (i) NP → noun *Bananas* are a good source of potassium.
- (ii) NP → det noun *My doctor* told me to exercise more.
- (iii) NP → adj noun *Easy melodies* make for good songs.
- (iv) NP → det adj noun *The old ways* are the best ways.
- (v) NP → det noun PP *The road to Hell* is paved with good intentions.

By positing these five different, but related, rules that expand *NP*, we are saying that any *NP*, any place in a sentence, can have any of those five structures. To repeat: that is not entirely true, but it is a good first step to take in approximating the way words are ‘distributed’ in English and in other languages.

It is often the case that we can simplify our analysis of a phrasal category by saying that a part of its expansion is *optional*. Instead of saying that we have both rules (i) and (ii) above, we say that *det* is optional, and the notation for that is a set of parentheses around the optional category:

(13) NP → (det) noun.

Looking at all of the expansions given in (12xx), we would naturally be led to the conclusion that a better form of the *NP* rule would be this:

(14) NP → (det) (adj) noun (PP)

(Discuss the consequences: more expansions predicted now.)

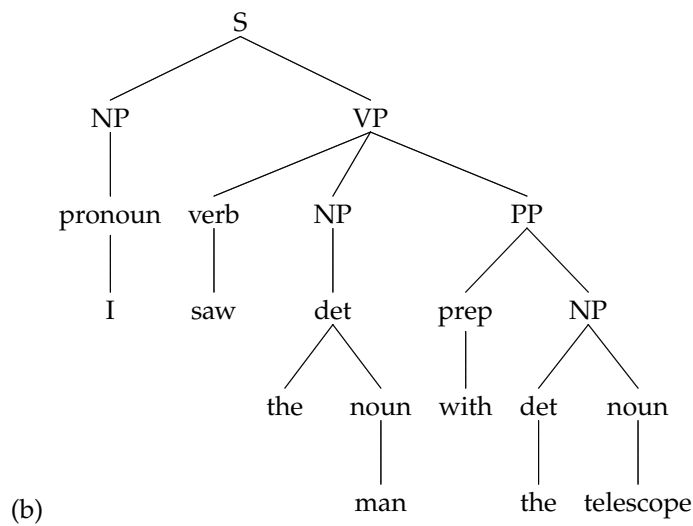
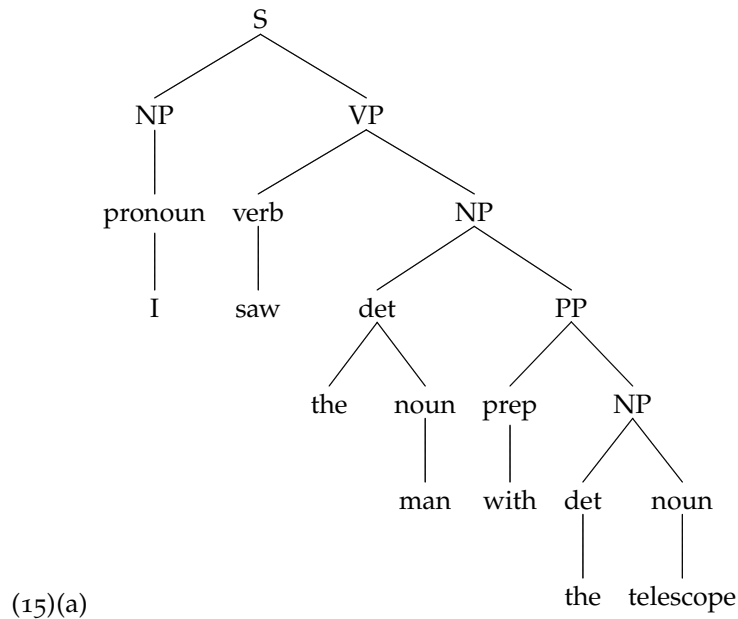
Ambiguous sentences

In analyzing ambiguous sentences, most of the time we assign two different syntactic structures, one with each of the intended interpretations, as we did with sentences (1a) and (1b), and in most of these cases, there are two or more words which are assigned different lexical categories in the two cases. In the sentence we considered, “Left” was a noun in the intended sense—perhaps a noun derived from a verb, but in any event, it referred to a political party, or a coalition of parties. In the unintended sense, “Left” was the main verb of the sentence, the past tense of the verb *leave*. Our analysis, then, predicts that if we change the word “Left” into some other word, some word that is not both a verb and a noun, the sentence should become unambiguous and not funny at all. That is true: there is no humor in *British Right Waffles on Falkland Islands*, or in *British Leave Waffles on Falkland Islands*. The humor of the ambiguity arises out of the totally unexpected collision between two different syntactic structures, themselves the result of simple phrase-structure rules motivated by an enormous number of simple rules.

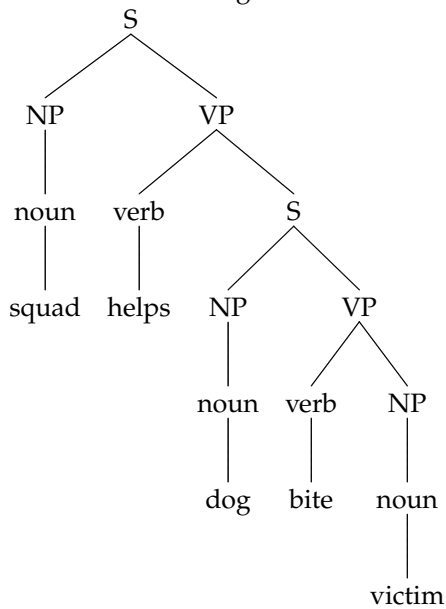
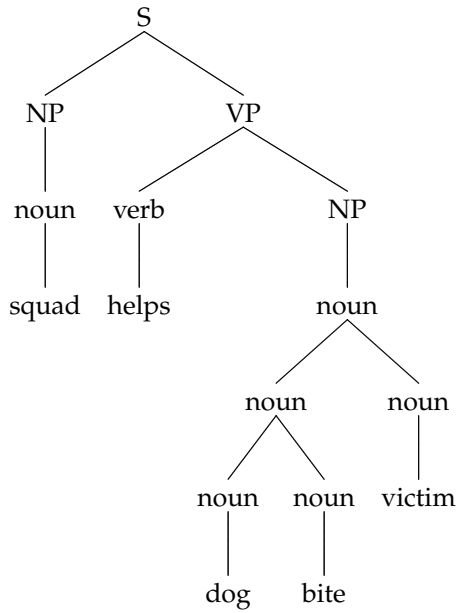
By the way: not all ambiguities are like that; one of the most over-used ambiguous sentences, *I saw the man with the telescope*, is ambiguous in a strictly structural way. Is it the man with the

We do not always know when an ambiguous sentence is syntactically ambiguous. Is *they are married* ambiguous? If not, where does the humor come from in *They're married, but not to each other*? How about *Kids make nutritious snacks*? That is ambiguous, but it may not be syntactically ambiguous. And what about *My father always beat me... at chess, at least*?

telescope that I claim to have seen, or am I just talking about some man and the fact that I looked at him through the telescope? These two senses correspond to two different syntactic structures:



Let's consider another ambiguous sentence:



The second structure arises unambiguously if we put in some words that allow no other analysis — for example, if the sentence had been *squad helps dog find master*.

Constituents

Any string of words that is generated by a single phrasal node in a given sentence is called a *constituent*. To analyze a sentence is to assign a tree structure to it, and by doing so, to analyze a set of constituents in the sentence. A good part of syntactic analysis is finding the right constituency structure for a sentence (we sometimes say, the right tree structure).

The most direct way to apply tests for constituency is to use the independence assumption that I mentioned earlier: if a string of words is a constituent – an NP, let’s say – then it ought to be possible to use that string of words in other sentences that seems

structurally rather different. If a string of words is a direct object NP (*the price of tea in Japan* in the sentence *we compute the price of tea in Japan*), then it ought to be possible to put the same string of words in places where we are already pretty sure that NPs can appear, such as in subject position of a simple sentence, or as the object of a preposition:

(16) The price of tea in Japan drives economic conditions there.

(17) I don't know much about the price of tea in Japan.

or other constructions, such as the pseudo-cleft:

(18) What they study is the price of tea in Japan.

or the cleft (formed with *it*):

(19) It was the price of tea in Japan that was the most important factor, not the temperature in Seattle.

What does this test suggest about the constituency of *The congregation sent the family flowers*? Is *the family flowers* a constituent? The fact that the following strings of words are not good sentences suggests strongly that it is not a constituent.

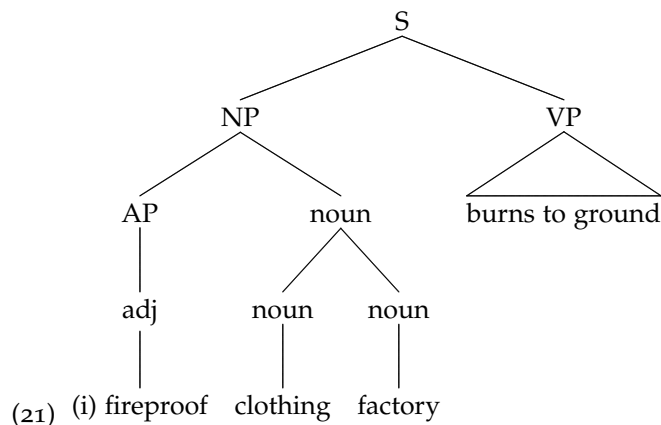
(20)(a) *What they sent was the family flowers.

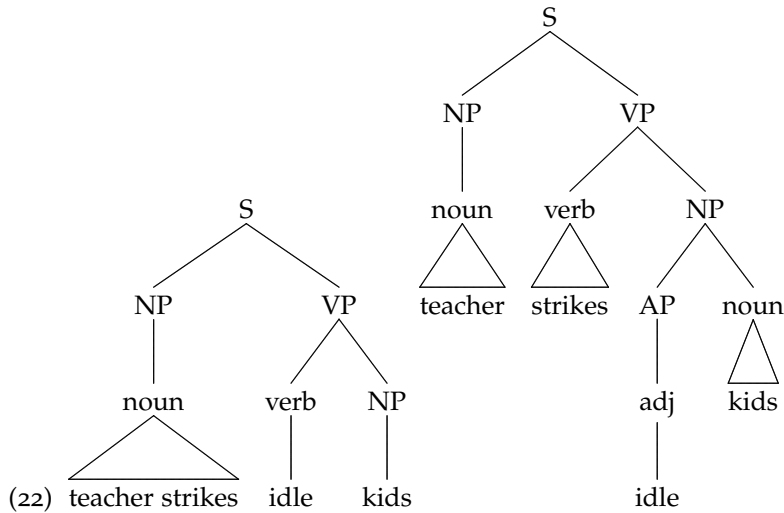
(b) *It was the family flowers that they sent.

We will look shortly at the difference between *John turned over the book* and *John jumped over the puddle*. Can you tell if *over the book* or *over the puddle* is a constituent?

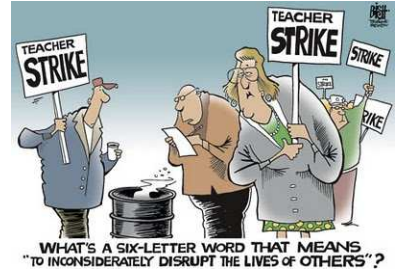
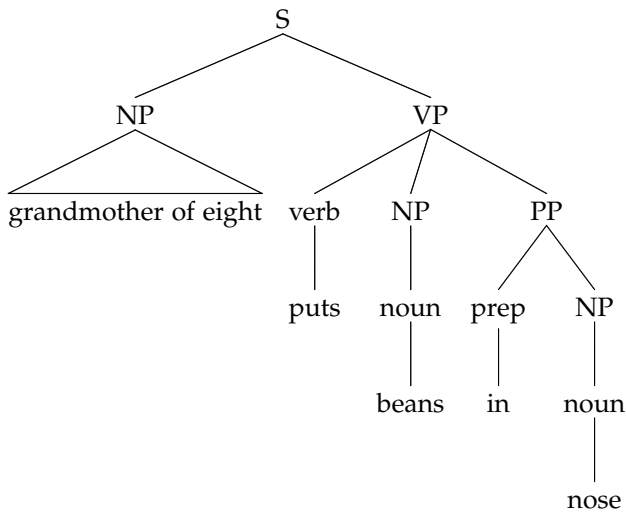
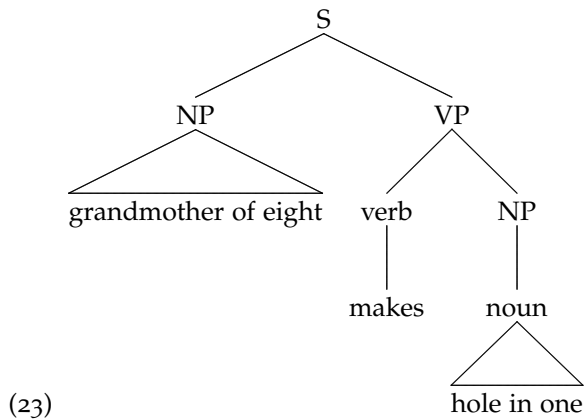
More examples

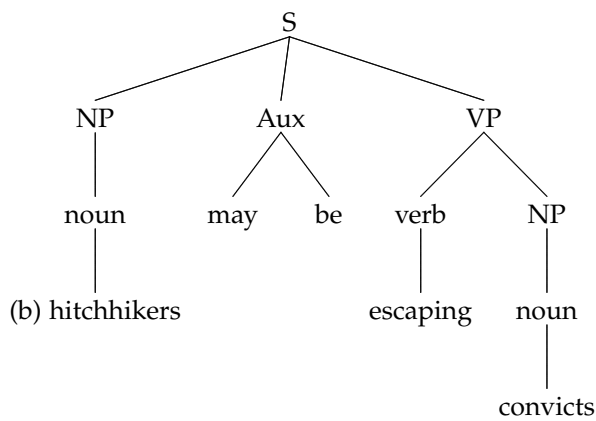
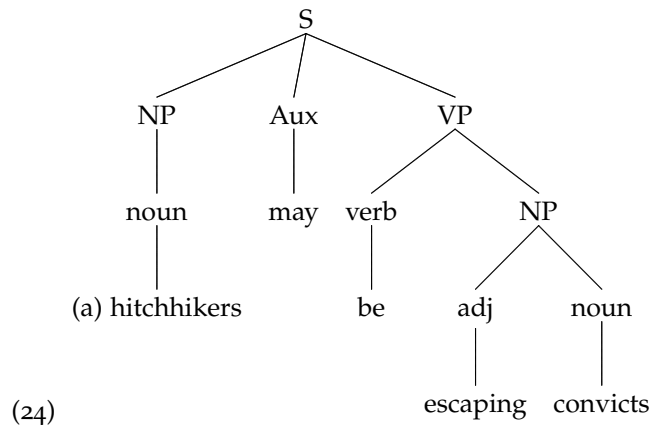
A simple example illustrating constituent structure ambiguity:
Fireproof clothing factory burns to ground.





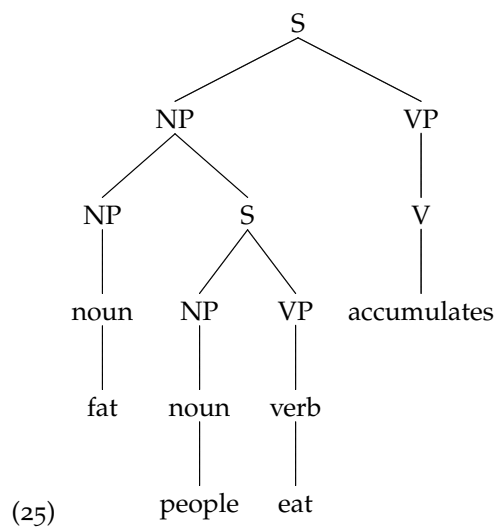
The interest of the headline: *GRANDMOTHER OF EIGHT MAKES HOLE IN ONE* relies on a structural difference: is [*hole in one*] a single item, or does it form two “sister constituents” in the verb phrase, as in *she put it in the bag* (or “...puts beans in nose”)?





Another nice way to sensitize oneself to syntactic structure is to look at garden-path sentences, like

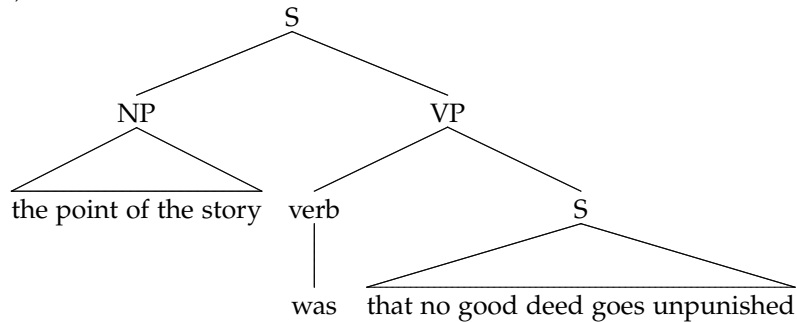
1. Fat people eat accumulates.
2. The cotton clothing is usually made of grows in Mississippi.
3. The girl told the story cried.
4. The horse raced past the barn fell.
5. I know the words to that song about the queen don't rhyme.



Infinitives and embedded clauses

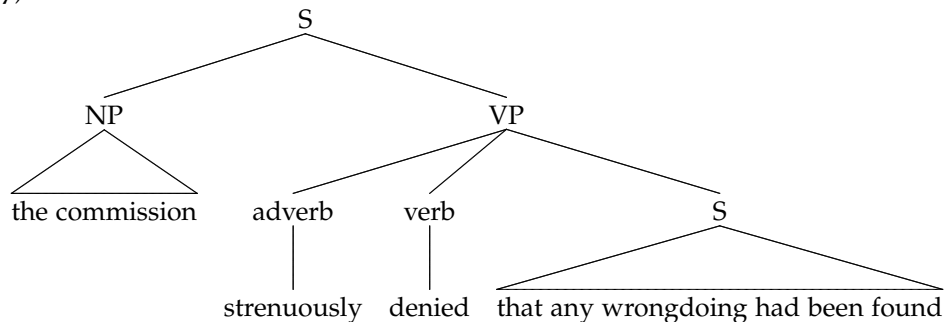
We generally use the term *clause* a bit more generally than the term *sentence*. We often find that what could be a free-standing sentence is part of—or, as we say, is *embedded in*—a larger clause. Consider:

(26)



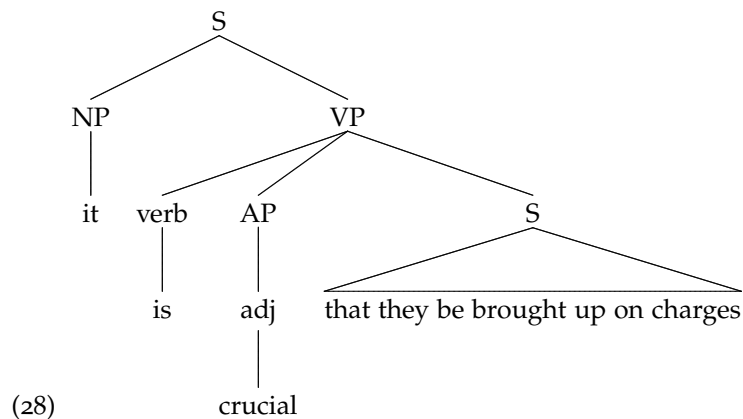
No good deed goes unpunished can appear as a free standing sentence, and it appears in (x) as an embedded clause. Sometimes an embedded clause has largely the structure of a free-standing clause, though some parts of it are affected by the sentence in which it is embedded, as in this example:

(27)

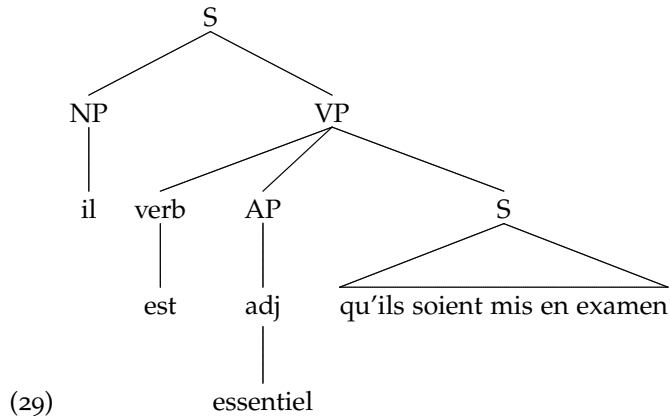


Any wrongdoing had been found cannot form a free-standing sentence: the possibility of the *any* in the embedded clause is the result of the negative sense that is implicit in the verb *denied*.

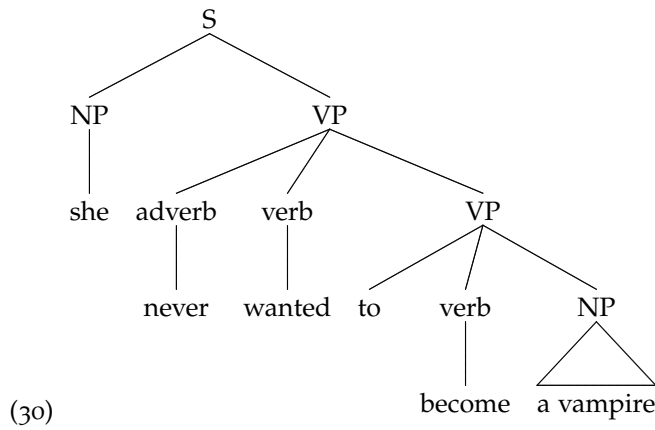
Thus embedded clauses may look different from main clauses. Sometimes the verb takes on a special form, as in the next sentence, or in a French sentence where the embedded clause has a verb in the subjunctive.



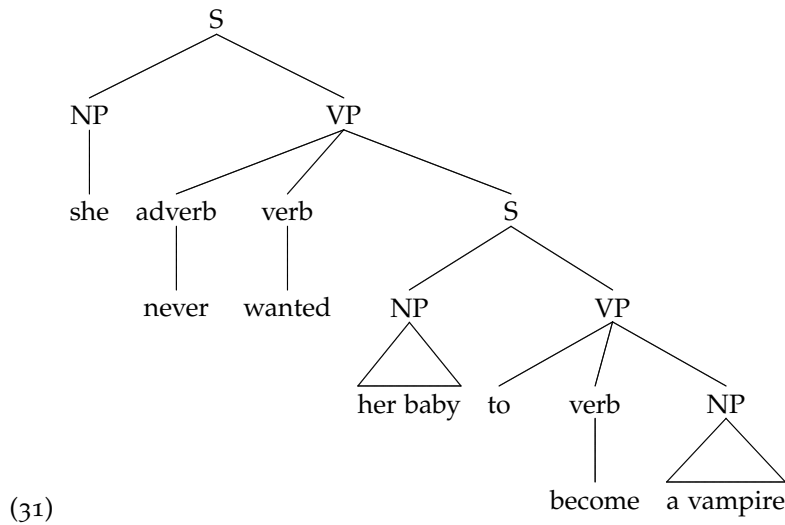
(28)

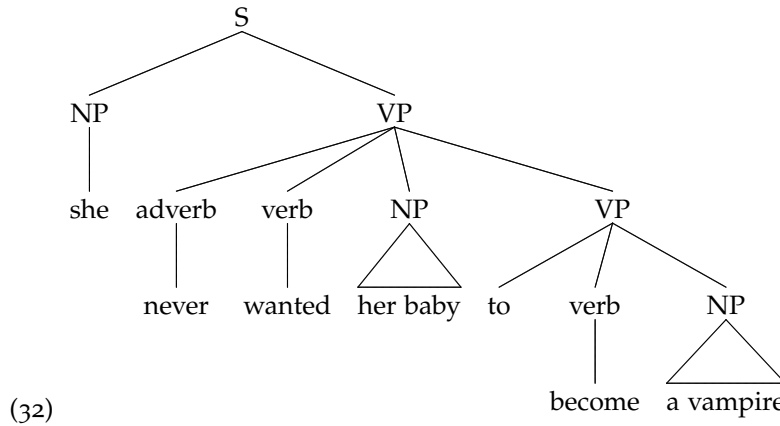


In many languages, the form of the embedded clause is considerably reduced when the subject of the embedded clause refers to the same person or think as the subject of the higher clause—we say, when the subject of the upper and the lower clauses co-refer, as in:

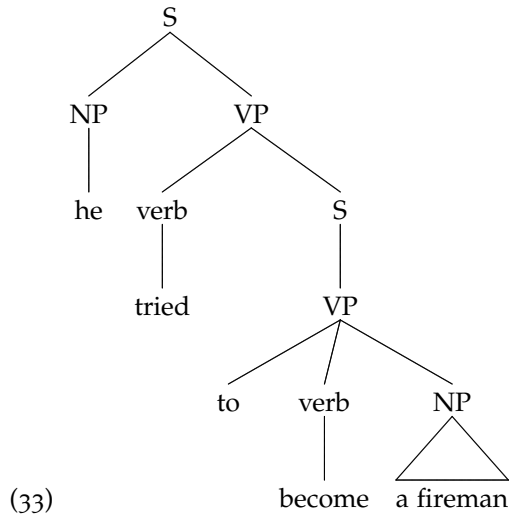


The embedded clause in that sentence could have a different subject, though it is a point of some controversy as to whether that sort of sentence – *She never wanted her baby to become a vampire*, for example — has the structure in (a) or in (b):





So: although there is controversy regarding the precise details of the analysis, let's agree to represent verb phrases with an infinitive as VP (verb phrases) immediately dominated by S:



Auxiliary verbs

One of the most impressive and influential of the early generative analyses of English was Chomsky's analysis of the English auxiliary. Let's consider a range of possible auxiliary verb combinations.

There is one thing that separates this data from the kind of data we have considered up to now. In the earlier examples, the choice of words that we made was essentially irrelevant; we included words by selecting nouns where the phrase structure rules generated "noun", and likewise for the other categories. But here – each word or morpheme acts differently and uniquely. Why would we expect phrase-structure rules to work here? Either we will have actual words in our phrase-structure rules, or we will have to create categories that contain only a single item. The two pretty much boil down to the same thing.

	You				walk.
	John				walk -s.
	John				walk-ed.
	John	may			walk.
	John	may	have		walk-ed.
	John		has		walk-ed.
	John			is	walk-ing.
	John	may		be	walk-ing.
	John	may	have	be-en	walk-ing.
Sentences with <i>-ed</i> :	John	may	have		walk-ed.
	John		has		walk-ed.
	John				walk-ed.
Sentences with <i>-ing</i> :	John			is	walk-ing.
	John	may		be	walk-ing.
	John	may	have	be-en	walk-ing.
Sentences with 3rd p. sg -s:	John				walk -s.
	John			is	walk-ing.
	John		has		walk-ed.
Sentences with <i>-do</i> :	You	<i>do</i>			walk.
	John	<i>does</i>			walk.
	*John	<i>does</i>			walk-s.
	*John	does may	have		walk-ed.
	*John	does	has/have		walk-ed.
	*John	does		is/be	walk-ing.
	*John	does may		be	walk-ing.
	*John	does may			walk.
	*John	do may	have	be-en	walk-ing.
Do	you				walk?
Does	John				walk?
May	John				walk?
May	John		have		walk-ed?
Has	John				walk-ed?
Is	John				walk-ing?
May	John			be	walk-ing?
May	John		have	be-en	walk-ing?
	*You		not		walk.
	You	do not			walk.
	*John		not		walk -s.
	John	does	not		walk.
	John	may not			walk.
	John	may not	have		walk-ed.
	John		has not		walk-ed.
	John			is not	walk-ing.
	John	may not		be	walk-ing.
	John	may not	have	be-en	walk-ing.

	You			were	amaze-d.	
	John			was	amaze-d.	
	John	may		be	amaze-d.	
	John	may	have	be-en	amaze-d.	
	John		has	be-en	amaze-d.	
	John			is	be-ing	amaze-d.
	John	may		be	be-ing	amaze-d.
	John	may	have	be-en	be-ing	amaze-d.
	You			were <i>not</i>	amaze-d.	
	John			was <i>not</i>	amaze-d.	
	John	may <i>not</i>		be	amaze-d.	
	John	may <i>not</i>	have	be-en	amaze-d.	
	John		has <i>not</i>	be-en	amaze-d.	
	John			is <i>not</i>	be-ing	amaze-d.
	John	may <i>not</i>		be	be-ing	amaze-d.
	John	may <i>not</i>	have	be-en	be-ing	amaze-d.

Table 1: English auxiliary

Let's try to extract some basic generalizations concerning this data:

- No sentence with two words from the group called **modal verbs**: *may, can, will, would, may, should, shall* is grammatical; but one word from this group can co-occur with the other auxiliary verbs, such as *have, be*.³
- When auxiliaries appear, their left to right order is summarized by a table:

Modal verb	have (perfective)	be (progressive)	be (passive)	verb
------------	-------------------	------------------	--------------	------

- The auxiliary verb *do* does not appear when there is any other auxiliary present: any of the auxiliaries we are exploring. It only appears when there are no others.
- However, the auxiliary *do* can appear along with the possessive *have* and the real (not dummy) verb *do*: *We do not have enough money to do that. Anyway, we do not do things like that.*
- If the negative *not* is present, it appears after the left-most (i.e., the first) of all of these auxiliaries. And if we count the auxiliary *do* as belonging to this group (and we do!), then when there is a *not*, there *must* be an auxiliary.

³ Well. Most of us know that this isn't really true. There are a lot of speakers of American English in the South who say *I might could give you a hand: might could*, and for many, *might could* and even *may can*. This analysis is *very* hard to modify to include those.

Chomsky and Syntactic Structure: the basics

Chomsky's account in *Syntactic Structures* (1957) was along the lines of what I have put in Figures 1 and 2 (I have made some changes that I think no one would disagree with, with hindsight).

Chomsky's example was more like the Figure 4. He alluded to morphophonemic rules that would include *will + S → will, will + past → would*.

$S \rightarrow NP Aux VP$
 $Aux \rightarrow Tense(Modal)(have + en)(be + ing)(be + en)$
 verb \rightarrow hit, take, walk, read, etc.
 modal \rightarrow will, can, may, shall, must
 $Tense \rightarrow S / NP_{sing} \text{ ---}$
 $Tense \rightarrow \emptyset / NP_{e'}$
 $Tense \rightarrow past$
Affix hopping:

Figure 1: English auxiliary (after Chomsky 1957)

$$\left\{ \begin{array}{l} past \\ S \\ \emptyset \\ -en \\ -ing \end{array} \right\} \left\{ \begin{array}{l} Modal \\ verb \\ have \\ be \end{array} \right\} : 1 - 2 \rightarrow 2 - 1\#$$

Chomsky suggests an abbreviation of *Af* for the disjunction

$$\left\{ \begin{array}{l} past \\ S \\ \emptyset \\ -en \\ -ing \end{array} \right\}.$$

Replace + by # except in the context v-Af.

Insert # initially and finally.

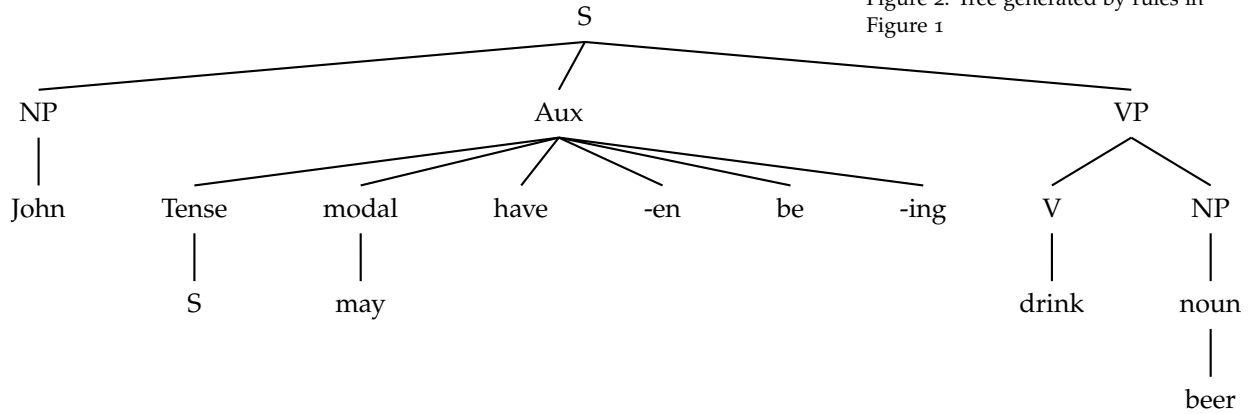


Figure 2: Tree generated by rules in Figure 1

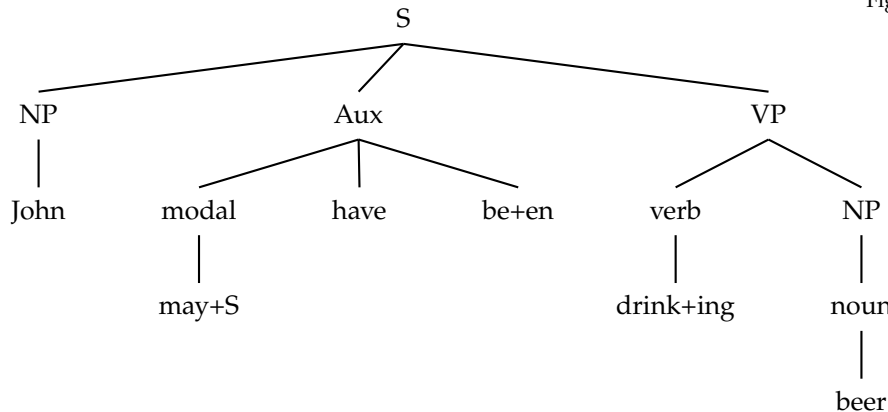


Figure 3: After affix-hopping

the	+ man	+ Aux		+ VP		
the	+ man	+ Aux		+ verb	+ NP	
the	+ man	+ Aux		+ verb	+ the	+ book
the	+ man	+ Tense	+ have + en + be+ing	+ read	+ the	+ book
the	+ man	+ S	+ have + en + be+ing	+ read	+ the	+ book
the	+ man		+ have + S # +be + en #	+ read + ing #	+ the	+ book
the	# man		# have + S # +be + en #	# read + ing #	# the	# book
the	man		has been	reading	the	book.

Chomsky's negation transformation

NP - Tense - X → NP - Tense + not + X

they -∅ + can + come	they -∅ + can + not + come
they -∅ + have -en + come	they -∅ + have + not -en + come
they -∅ + be-ing + come	they ∅ + be + not -ing + come
John - S - come	John - S + not - come

Affix hopping applies *after* the negation-insertion transformation, and cannot apply, because the *not*, like a grain of sand in the gears, prevents the rule from finding the context it is looking for. Chomsky adds a *later* rule (known to all later on as *do*-Support), which applies after all of the rules mentioned above:

(34) *Do*-support: # Af → #do + Af

Shortly after this (p. 65), Chomsky proposes a transformational rule that introduces a morpheme called *A* whose realization is as emphasis on the word that precedes it. In this case, the appearance of a form of *do* when there is emphasis ("John *does* arrive") is accounted for by the linear placement of *A* that is (i) in the same spot as the *not*, and (ii) equally able to block the hopping of the *S*-affix; which failure to hopping leads to an *S* which triggers *Do*-support. Imagine a derivation containing the step: *John # S+A # arrive*, and you have it.

See Figure 4 for a slightly different constituency structure.

Constituents -2

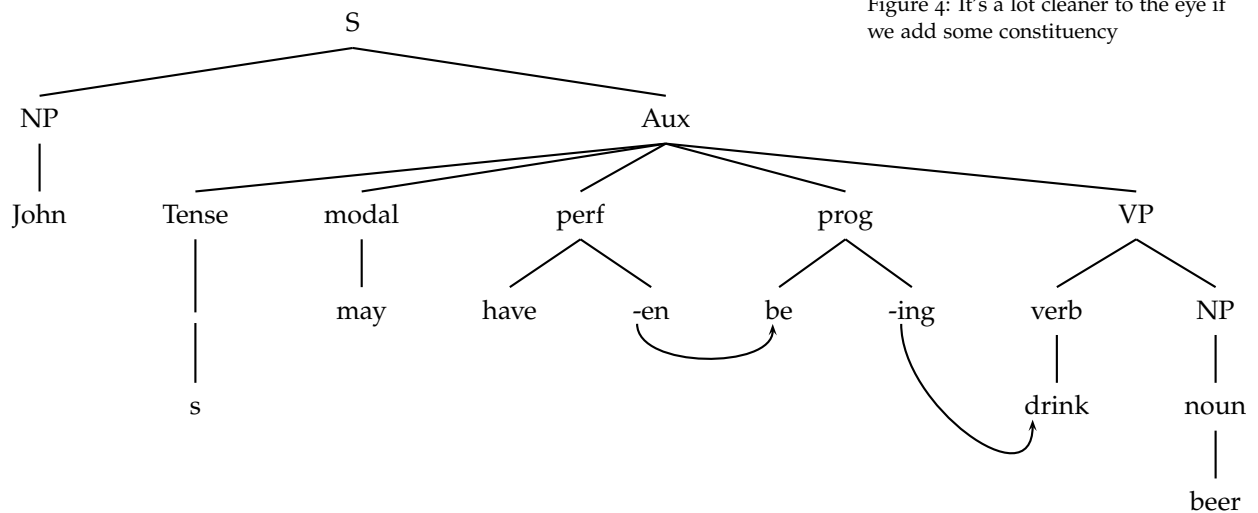
NP Verb PP; NP Verb NP PP

Our first look at some of the details of English syntax involved the auxiliary verbs. A very different kind of syntactic distribution is found when we look at what *follows* the verb in English. There are, to be sure, many *intransitive* verbs in English, as in (xx), where nothing follows the verb. There are also many in which a noun phrase follows the verb – we call these *transitive* sentences, as in (xx) – as well as many which are followed simply by a prepositional phrase (xx).

Peacock was born to hustle, bustle, jostle, and command, but he had as well a clear-eyed sense of who in the English mathematical establishment could be counted on, who counted in, and who counted out. David Berlinsky, *One, Two Three*. p. 93.

"How many people work at your company?"

"About half..."



Aux → *Tense*(*Modal*)(*have* + *en*)(*be* + *ing*)(*be* + *en*)

perf → *have* + *en*

prog → *be* + *ing*

passive → *be* + *en*

- (35) 1.(a) The baby is sleeping.
 (b) Whenever it rains, it pours.
 (c) Man plans, and God laughs.
- 2.(a) I love salmon, but Jessie can't eat it.
 (b) The contractor has finished the kitchen.
 (c) The House finally passed the president's legislation.
- 3.(a) All rivers run to the sea.
 (b) She spoke to every expert she could find.
 (c) Dr. King dreamt of a world in which all men are brothers.
 (d) Do not speak to the driver while the vehicle is in motion.

And finally, there are many sentences in which the verb is followed by a noun phrase and a prepositional phrase (see (37)).

(36) She put her name on the door.

(37) I translated the text into French.

In class we discussed some of the basic heuristics for getting information about constituency, such as:

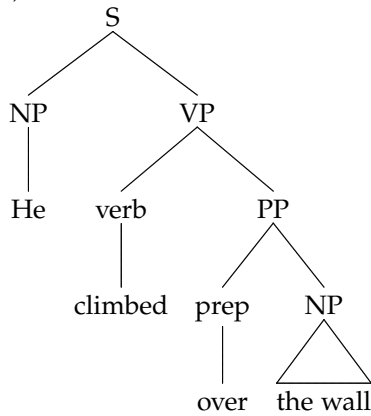
1. We can look at constructions which select a single constituent in a given position (subject of a sentence; focus of (*it*)-cleft, focus of pseudo-cleft), and see what string of words can show up in those positions;
2. if we can replace a string of words by *it* and retain the syntactic construction, this suggests the string is an NP;

3. if we can coordinate two strings with *and*, this suggests that each is a constituent, and that together they form a constituent.

The syntactic patterns *NP Verb PP* and *NP Verb NP PP* are very common patterns in English and other languages. Let's take a look at several patterns of this general sort:

He climbed over the wall

(38)



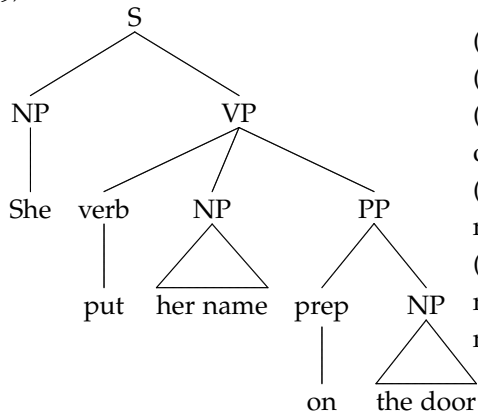
- (a) What did he climb over?
- (b) Over what did he climb? (maybe)
- (c) Over the wall climbed the monkeys.
- (d) Over the wall the monkeys climbed. (maybe)
- (e) The wall was climbed over. (maybe)
- (f) This wall has never been climbed over.
- (g) He climbed over it.
- (h) He climbed over the wall and the hedges.
- (i) He climbed over the wall and through the thick brush on the ground.

The (b) example—if it is grammatical—is evidence that *over* and its following object VP forms a constituent; in the metaphor of syntactic movement, a preposition would only move with its object. (c) (which is, I think, unquestionably grammatical) makes the same point, but in the context of a different construction. (e) is a passive, in which the object of *over* has been passivized; this suggests a tight syntactic relationship between *over* and the preceding verb *climb*, and if (e) is not great, (f) is, and it makes the same point regarding grammar.⁴

⁴ The point is often made in relation to the contrast between *This bed has been slept in* and *This bed has been slept under*, where the first is much better than the second.

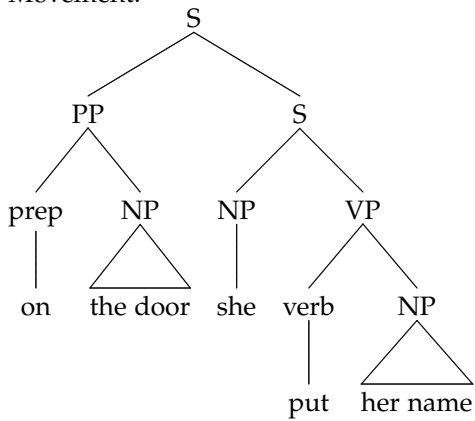
She put her name on the door

(39)

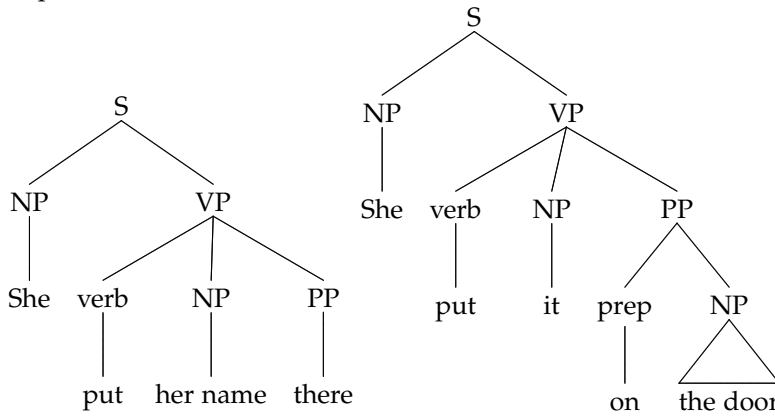


- (a) What did she put on the door?
- (b) Where did she put her name?
- (c) What did she put her name on?
- (d) On the door, she put her name.
- (e) On the door, she put her name; on her desk, she put her new title.

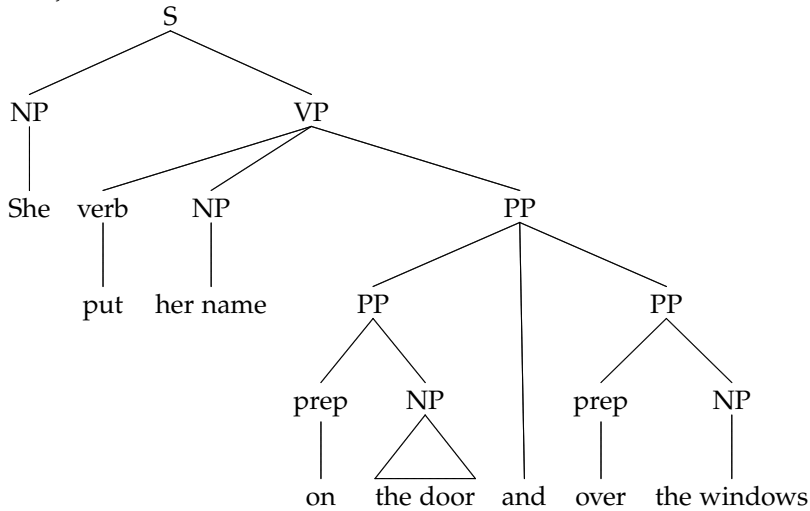
Movement:



Expansion:



Conjunction:



They turned out/off the light

Now, let's consider the sentence *They turned out the light*, which is also of the form NP V P NP. Does this have the same structure? – that is, is it:

The first sign that this is *not* the same structure is that this structure is unavailable when we have *it* rather than *the light* (remember, this was fine with *he climbed over it*):

(40) 1. *They turned out/off it.

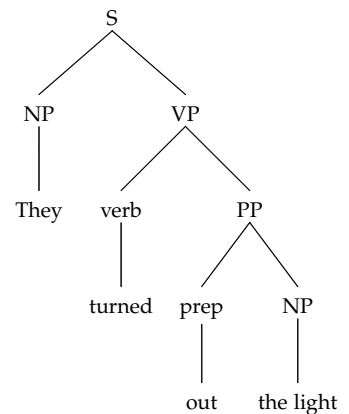


Figure 5: Wrong analysis!

2. They turned it out/off.

It is odd that *the light* cannot be simply replaced by *it* in *They turned out the light*, especially since apparently similar sentences are fine. Is this phenomenon general, fairly general, or just marginal? How can we check? Are there words other than *out* that participate in this oddity?

This is known as a *verb particle construction*, or as a *phrasal verb*.

to turn on something

(41) The lion turned on his trainer, and it was several minutes before he could be removed from the cage.

(42) (Not: ...turned his trainer on...)

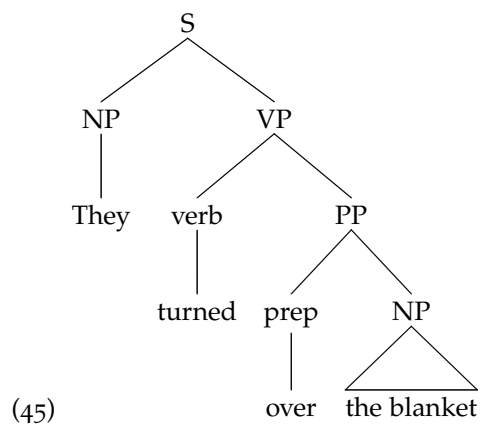
(43) The detective turn on her radio, and it was several minutes before she could tear herself away from what she was hearing.

(44) (just as fine...The detective turned her radio on...)

Questions: Do we wish to assign different structures to these sentences, and if so, how? What do you notice about the stress or prominence of the word *on* in the two sentences?

They turned over the blanket.

Is this right?



We can still say:

(46) What did they turn over?

but not:

(47) *Over what did they turn?

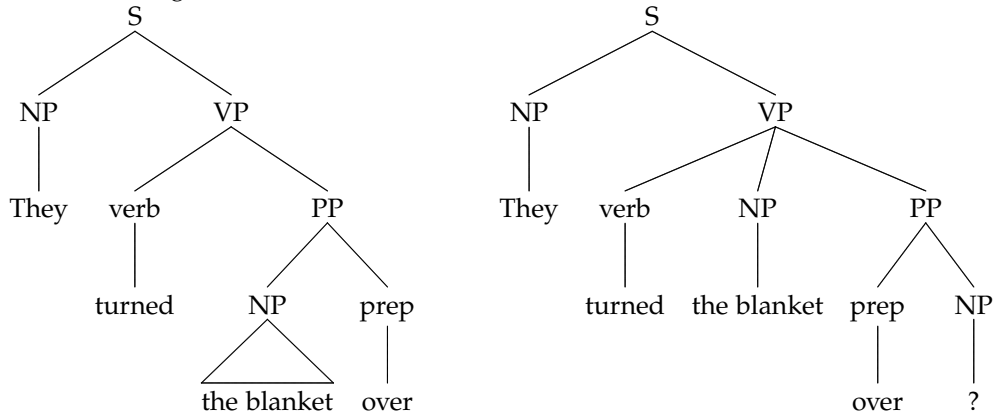
or

(48) *It was over the blanket that they turned.

So there is no evidence of *pied-piping*, of the preposition 'moving' along with the following NP. So *Over the blanket* does not behave like a constituent. And we can say:

(49) They turned the blanket over.

What is the right structure for *that* sentence?



What do we find if the object is a pronoun?

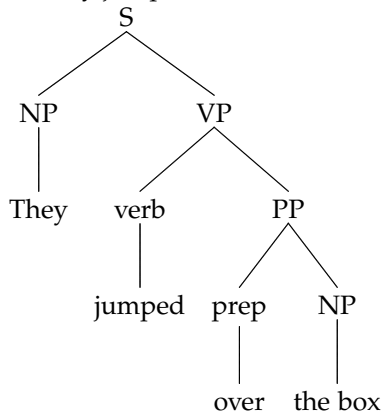
5

- (50) • They turned it/him over.
 • *They turned over it.

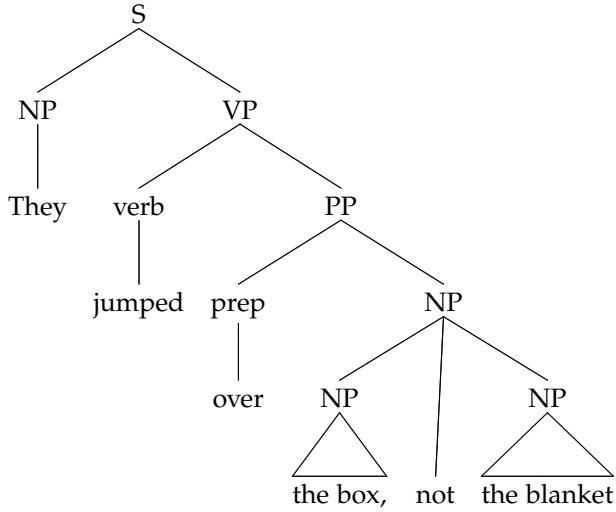
⁵ These facts might remind us of the similar ungrammaticality of **They gave Mary it*, alongside of the fine *They gave Mary some*.

They rolled it over/they rolled over it.

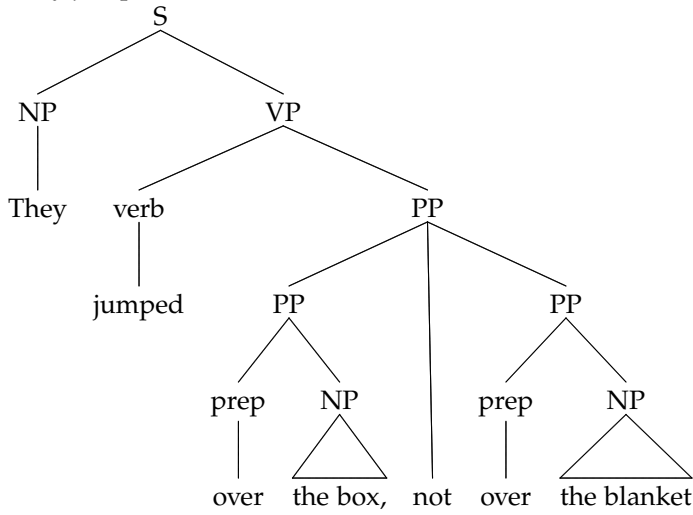
(51)(a) They jumped over the box.



(b) They jumped over the box, not the blanket.



(c) They jumped over the box, not over the blanket.

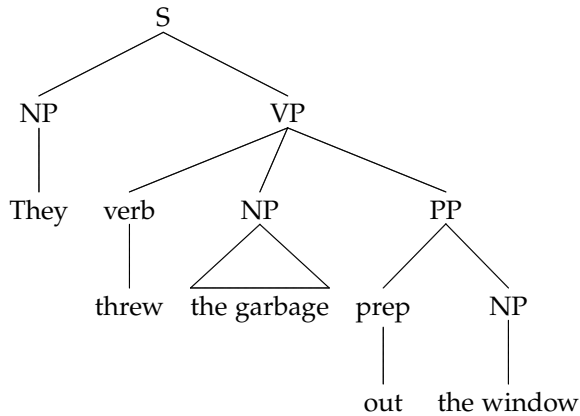


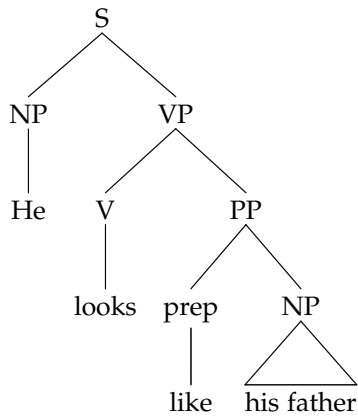
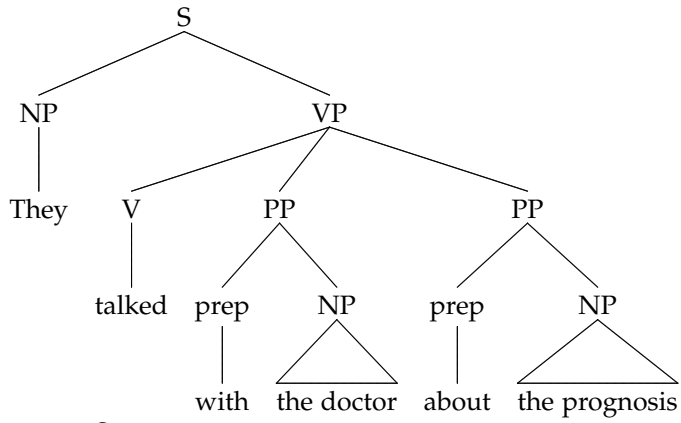
(d) They turned over the box.

(e) They turned over the box, not the blanket.

(f) **They turned over the box, not over the blanket.

They threw the garbage out the window.





- (52)(a) They jumped over the box.
- (b) They turned over the box.
- (c) They jumped over the box, not over the the shoes.
- (d) **They turned over the box, not over the shoes.
- (e) They turned over the box, not the shoes.

put	the book	on the table	
put	it	under the tree	
put	it	over the sink	
<hr/>			
put	the coat	on.	
put	the coat	on the monkey	
put	it	on.	
put		on	the coat.
put		on	*it.
put		on	shorts.
put		*on the monkey	the coat.
<hr/>			
put	the decision	off.	
put	it	off.	
put		off	the decision.
put		off	*it.
<hr/>			
take	the coat	off.	
take	the coat	off the monkey.	
take	it	off.	
take	it	off the monkey.	
take		off	the coat.
take		*off the monkey	the coat.
<hr/>			
drink	the water.		
drink	the water	(all) up	
drink		up	the water
drink		*all up	the water
drink	it	up.	
drink		up	*it.
drink	the water	out of the bottle	
?* drink	the water	up	out of the bottle.

What's the generalization? The direct object and the particle can permute—appear in either order—only if the particle is not part of a larger Prepositional Phrase. It cannot have a preceding determiner, and it certainly cannot have a complement (like *the monkey*).

Let's find some examples with *off*, *up*, *out*. Can we find any with *after*? *to*? *from*?

Some analyses

Thanks to Bas Aarts, "Verb-preposition constructions and small clauses in English" *Journal of Linguistics* 25(2): 277-290, 1989.

(53) **A-verbs** I switched the light off. (The lights are now off.)

(54) **B-verbs** I looked the information up. (The information is not now up, whatever that might mean.)

(55) A-verbs:

1. He propped the hood of the car up; with the hood up he then drove off.

2. Sally pushed the lever on the amplifier down; with the lever down her CD-player was pre-programmed.
3. Jim turned the radio off; with the radio off he could finally relax.

(56) B-verbs:

1. *He brought the kids up by himself; with the kids up he could go on holiday.
2. *My teacher always puts his pupils down; with his pupils down he feels superior.
3. *Jim sold the car off to a friend (now a former friend); with the car off he could buy the boat he had dreamed of.

(57) In comparatives, A-verbs are pretty good:

(58) A-verbs:

1. The oven off is less dangerous than the oven on.
2. The oven off is as dangerous as the oven on.
3. The ovens off is at least as dangerous as the ovens on. (What does this show?)

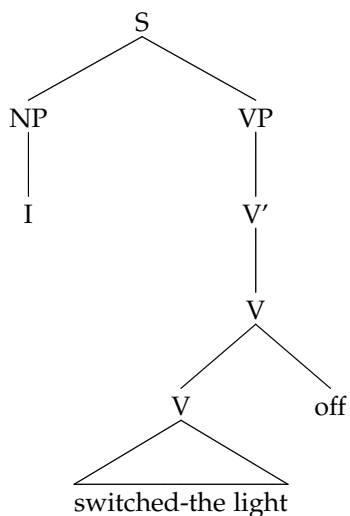
(59) B-verbs:

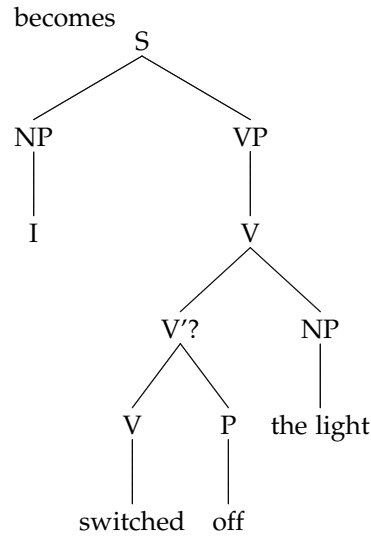
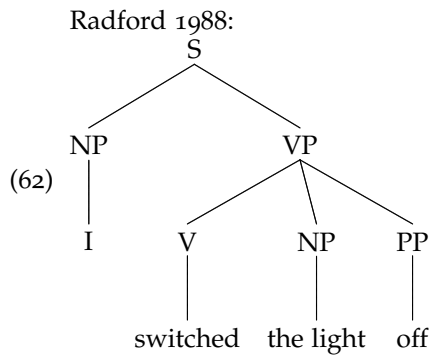
1. *He brought his kids up more than he brought them down.
2. *The kids up is very desirable.
3. *His pupils down is terrible (a terrible sight to behold).

(60) Conjunction: what does this show?

1. He switched the lights on and the TV off.
2. Compare: I gave Vincent a book and Caroline a newspaper.

(61) Stowell 1981:

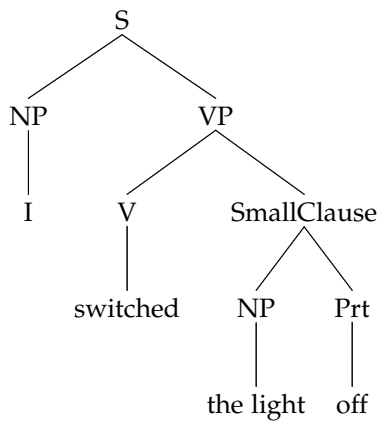




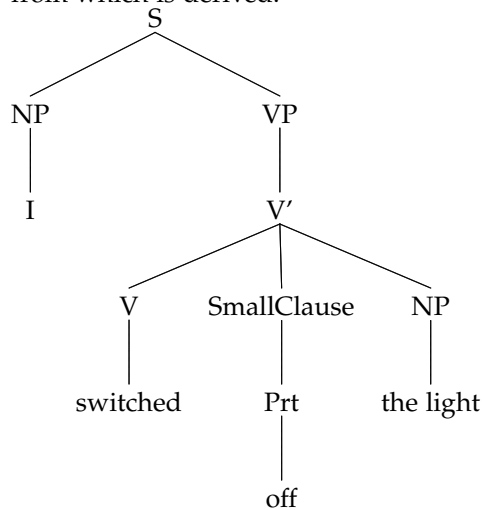
- (63) 1. I cut the branch right off.
 2. *I cut right off the branch.
 3. I switched the radio completely off.
 4. *I switched completely off the radio.

What do these show?
 That *off* is a phrase, not a single word – in the case where it is to the right of the direct object NP?

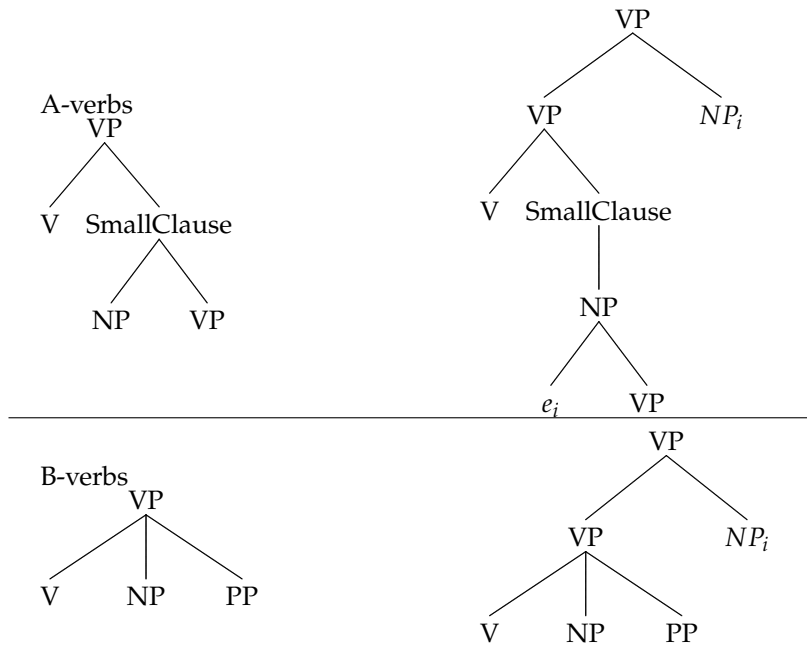
(64) Kayne 1984:



from which is derived:



(65) Aarts's analysis of A-verbs, B-verbs:



Some of the basic phenomena of interest to syntacticians

Word-order interacts with logical scope of operators

For example, in English:

Liberman 1975

- i. *With no job, John would be happy.* If he had no job (= if he were unemployed), John would be happy.
- ii. *With no job would John be happy.* There is no job such that it would make John happy (if it were given to him).

Basic word order: SVO and its permutations

Joseph Greenberg in 1966 drew attention to the fact that the order of constituents in sentences was not uniformly distributed among all the logical possibilities. Focusing on subject (S), object (O), and verb (V), studies (such as Ruhlen 1975) have found distributions along these lines:

www.hku.hk/linguist

SOV	SVO	VSO	VOS	OVS	OSV
52%	36%	10%	2%	0%	0.2%
VOS: Malagasy, Seediq (Austronesian)					
OSV: Kabardian (Northwest Caucasian)					
OVS: Apalai, Hixkaryana (Carib)					

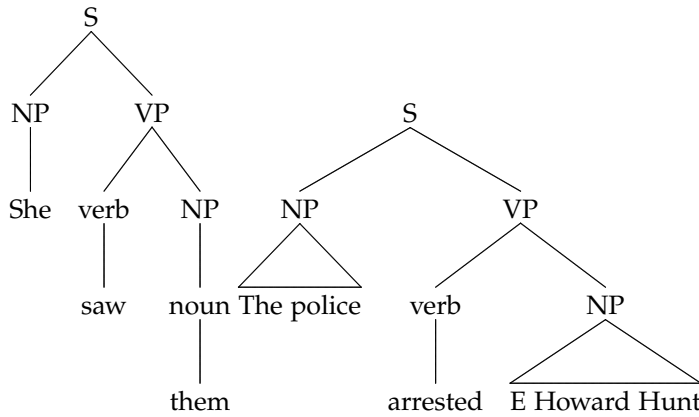
Pullum 1981

English: SVO

Subject-Verb-Object

The police arrested E. Howard Hunt.

S=sentence, NP = Noun Phrase, VP = Verb Phrase



Japanese: SOV

Japanese is a strictly verb-final language, with massive pro-drop and topic-marking (-*wa*). This combination is of great interest to many linguists.

Tanaka-san wa ringo -o tabemasu
 Mr. Tanaka TOPIC apple DO eat
 Mr. Tanaka eats the apple.

The preceding sentence would be a reasonable answer to the question: What does Tanaka-san eat? To answer, Who eats the apple?, you might say:

ringo -wa Tanaka-san ga tabemasu
 apple TOPIC Mr. Tanaka SUBJ eat
 Mr. Tanaka eats the apple.

Consider:⁶

⁶ from nihongo.anthonet.com

Tanaka-san ga kono ie ni sunde imasu.
 Mr. Tanaka SUBJ this house in living is.
 Mr. Tanaka is staying in this house.

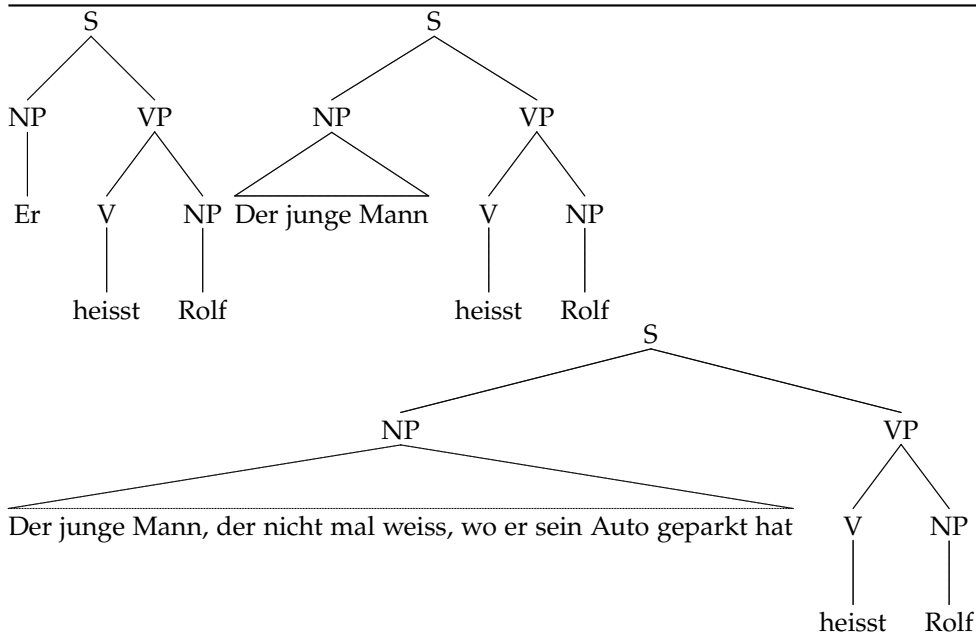
Tanaka-san wa sensei desu.
 Tanaka TOPIC teacher is.
 Tanaka is a teacher.

sunde ← sum+te.

German: mixed SVO, SOV

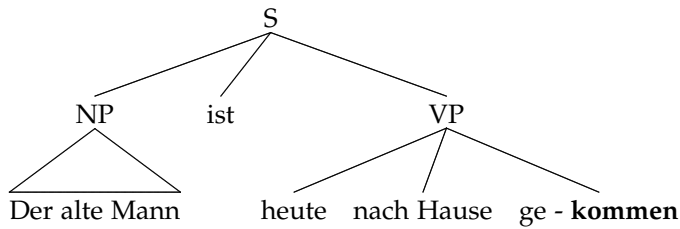
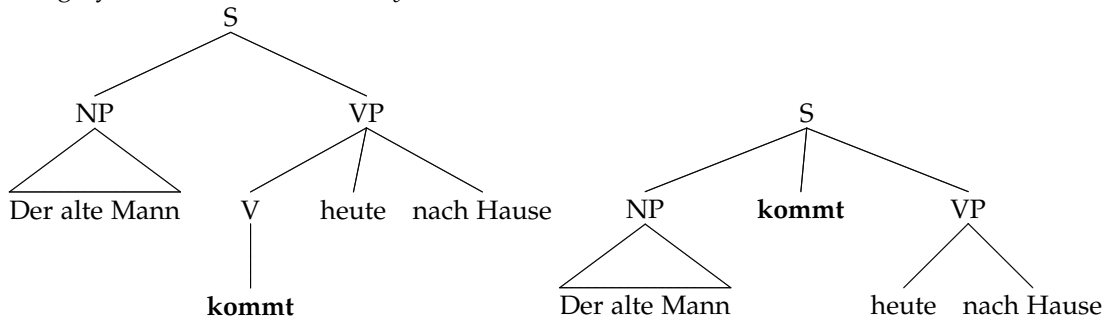
First approximation: In main clauses, the finite verb appears in second position, and a major syntactic constituent precedes it. A separable prefix does not appear in second position, even it is lexically associated with the verb that *is* in second position. When a series of verbs occurs in a single clause, the logically highest one is that which appears in second position. None of this occurs in embedded clauses – or rather, in sentences with overt complementizers.

	Wir	singen	Lieder.		
	Er	zieht	seinen Mantel	an.	
	Er	hat	seinen Mantel	an-ge-zogen.	
	Er	muss	seinen Mantel	an-ziehen.	
...	weil wir		Lieder	singen.	
...	weil er		seinen Mantel	an-zieht.	
...	weil er		seinen Mantel	an-ge-zogen	hat.
...	weil er		seinen Mantel	an-ziehen	muss.
	Sie	gibt	seinem Freund einen Apfel.		
	Sie	hat	seinem Freund einen Apfel	ge-geben.	
	Sie	hat	ihm einen Apfel	ge-geben.	
	Sie	hat	es ihm	ge-geben.	
...	weil sie		es ihm	ge-geben	hat.
	Sie	muss	es ihm	ge-geben	haben.
...	weil sie		es ihm	ge-geben haben muss.	
	Er	tut	es.		
	Er	hat	es	ge-tan.	
	Er	muss	es	tun.	
	Er	hat	es	tun müssen.	
...	weil er		es	hat tun müssen.	
	Ich	habe	Casablanca (nicht)	ge-sehen.	
	Casablanca	habe	ich (nicht)	ge-sehen.	
	Nie	hat	er Casablanca	ge-sehen!	
	Er	hat	mit dem Karl	ge-sprochen.	
	Mit dem Karl	hat	er	ge-sprochen.	
	Was	singen	sie?		
	Wer	singt	Lieder?		
	Was	hat	sie	ge-gessen?	
	Wen	hat	sie	ge-sehen?	
	Mit wem	hat	sie	ge-sprochen?	



[ex from www.dartmouth.edu/~german/]

Roughly: *The old man comes today home.*



Der alte Mann ist gestern ange**kommen**. (66)

Der alte Mann will heute nach Hause **kommen**. (67)

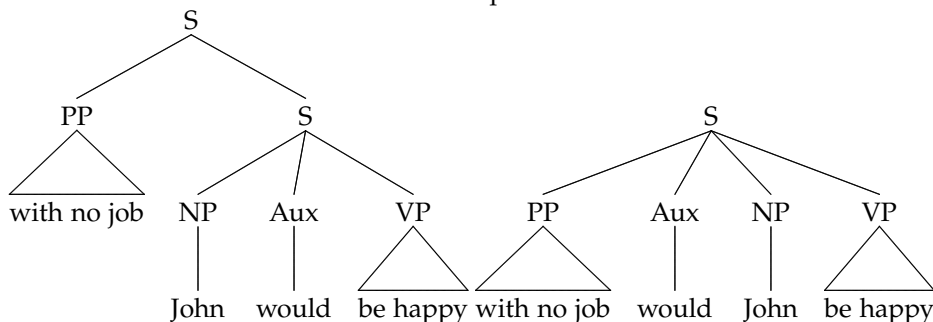
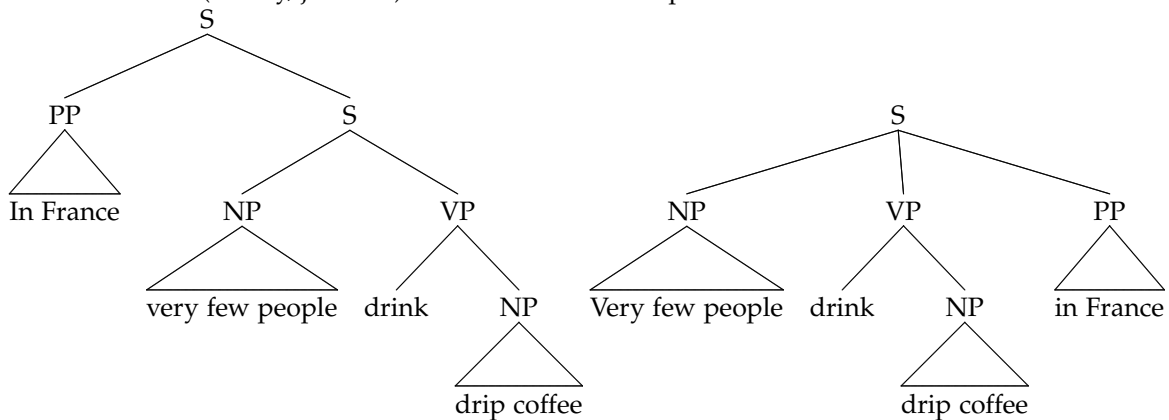
Heute **kommt** der alte Mann nach Hause. (68)

Ich weiss nicht, wann er heute ankommt. (69)

ex from german.about.com

There are a large number of phenomena that have been analyzed in terms of *syntactic movement*. *Movement* is, of course, a metaphor, but we use it to suggest a phenomenon whereby we have a good linguistic reason to analyze a word (or a constituent) as appearing in a position different from where it is on the surface.

Connection between constituent structure and movement: When we discover two closely related sentence patterns, we usually find that the difference can be expressed as a difference in the location of a small number (ideally, just one) constituent. For example:

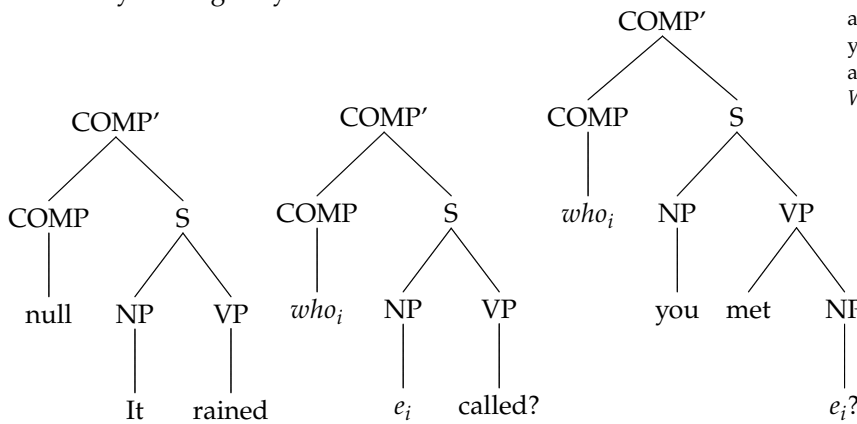


With no job would John be happy.

The clearest examples of this are the cases of question formation and, in many languages, relative clause formation.

Question formation

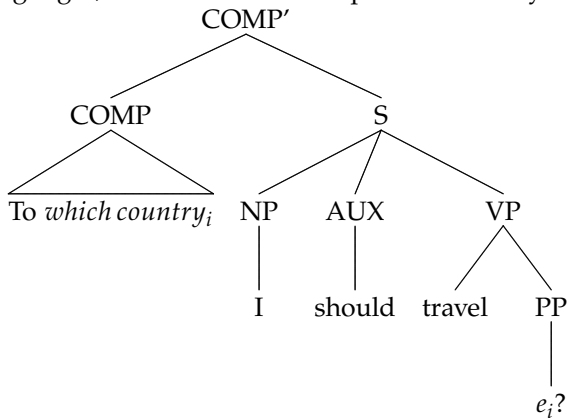
In English, a **question word** (or *wh-word*, or *whord*) appears sentence-initially in direction questions, even if it corresponds (in terms of the predicate of which it is an argument) to a NP in a different position. We will call the position in which *wh*-words are found the *complementizer* (or *Comp*) of a sentence. COMP' is read "COMP-bar", and is a shorthand for speaking of a larger constituent for which COMP is an obligatory member (even if it does not seem that the COMP really is obligatory here!).



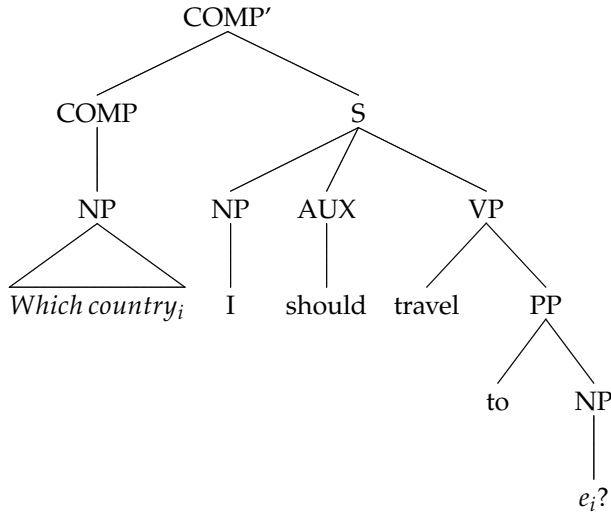
I leave the obligatory matter of subject-auxiliary inversion unstated here: but you should read the tree as if it had applied. The last example surfaces as *Who did you meet?*

This is called Pied-Piping

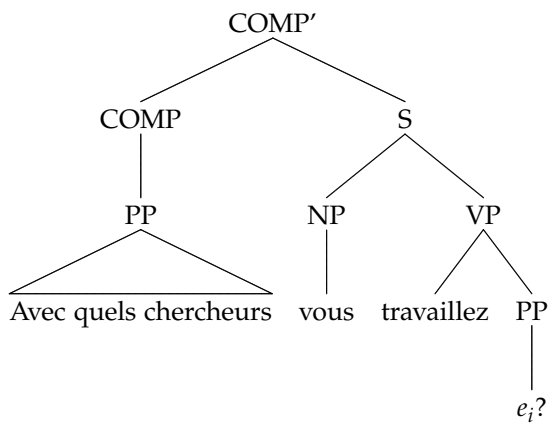
In *formal* English, a preposition may metaphorically move along with a *wh*-word, even if the preposition is part of an idiom along with the verb; while this is restricted to formal English, it is the normal and everyday case for many languages, include Romance languages; see the French example immediately below.



To which country should I travel?

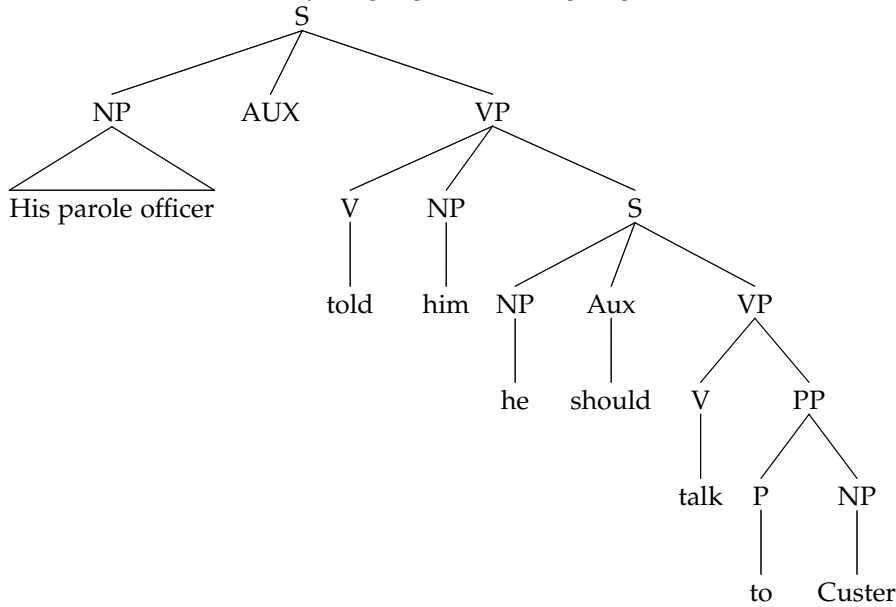


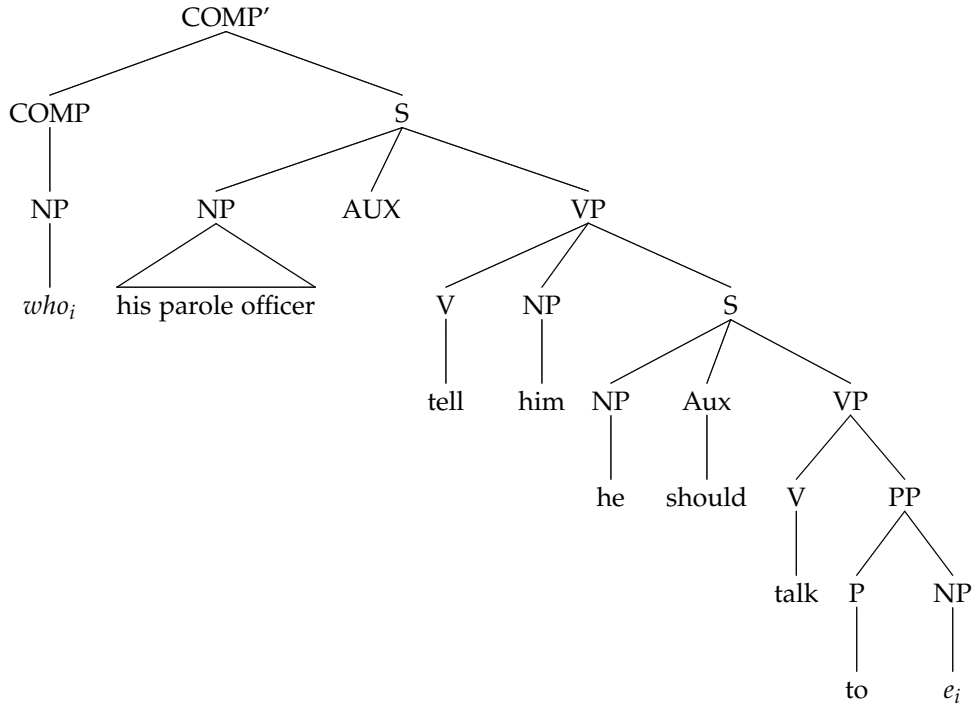
Which country should I travel to?



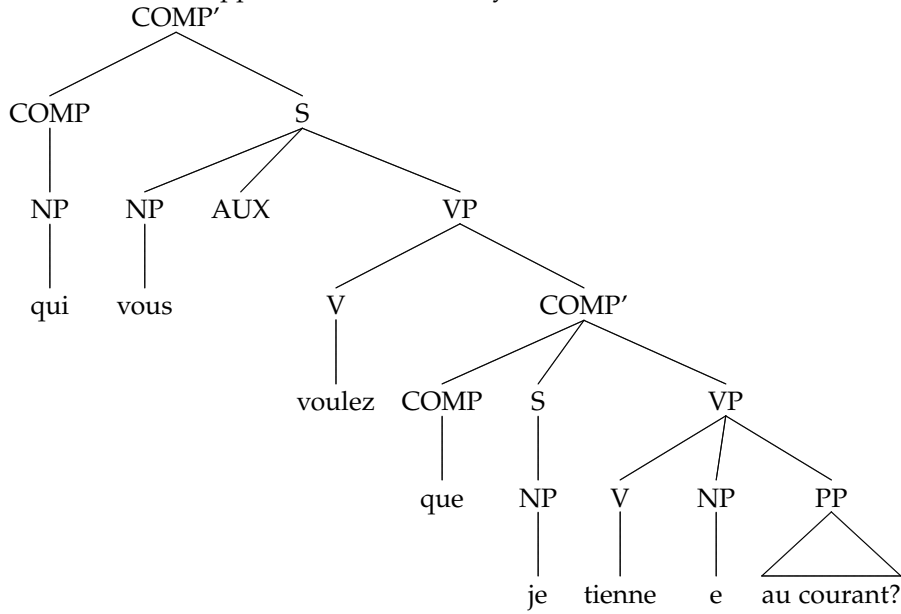
Here too I abstract away from inversion: cf. *Avec quels chercheurs travaillez-vous?*

With which researchers do you work?
 This wh-movement involved in question-formation can apply over several clauses, in many languages (including English).





In French, we see the verb of the main clause impose the subjunctive mood on the verb of the embedded clause, and the object of the lower clause appears sentence initially.



Who do you want me to keep [e] informed?

Qui voulez-vous je tienne [subj.] au courant?

Relative clauses

In English, a relative clause follows the head noun, and has a gap in the sentence corresponding to the position in which the head would have appeared in the relative clause:

the fruit $\left\{ \begin{array}{l} \textit{which} \\ \textit{that} \\ \emptyset \end{array} \right\}$ she had picked.

In relativizing from subject position, an empty COMP is not allowed:

We purchased some fruit $\left\{ \begin{array}{l} \textit{which} \\ \textit{who} \\ \textit{that} \\ \emptyset \end{array} \right\}$ was not ripe, unfortunately.

The words *which* and *who* are *wh*-words (*who* is for people, *which* for non-humans), and are analyzed as involving movement: pied-piping is permitted in these cases, but that is a complementizer, and there is no overt movement when it is present:

the people $\left\{ \begin{array}{l} \textit{with whom} \\ \textit{with which} \\ \textit{with that} \\ \textit{with}\emptyset \end{array} \right\}$ she had consulted were enthusiastic.

Relativization over a long syntactic distance is possible, just as with *wh*-questions:

The so-called magic bullet was the bullet [that [the Warren Commission argued [Oswald had used [e] to shoot both Kennedy and Connally.]]]

Question formation brings a *wh*-word to sentence-initial (COMP) position, but it can be a position at the beginning of a subordinate clause:

It was never determined what the former CIA employees were actually looking for [e] at the Watergate.

*What was it never determined the former CIA employees were actually looking for [e] at the Watergate?

Whose is both a relative pronoun and a *wh*-word, but it is specifically for humans as a *wh*-word, but not as a relative pronoun:

The *car_i whose_i door* was smashed in the accident had to be junked afterwards.

Whose_i door_i was smashed in the accident? OK: *Mary_i/mine_i*; **Mary's car's/that car's*.