

## Lecture notes: Week 8

### The complementizer phrase

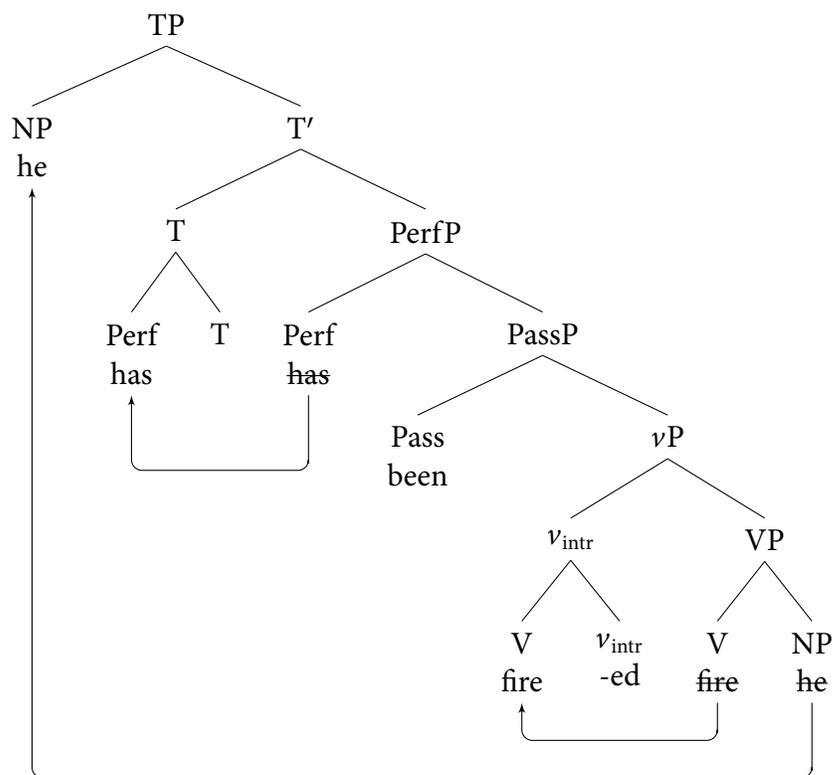
#### 1 Complementizers and subjects of embedded clauses

- So far, we have seen how to analyze complex sentences containing multiple auxiliaries:

(1) He has been fired.

- The structure we have proposed (omitting all features) is given below.

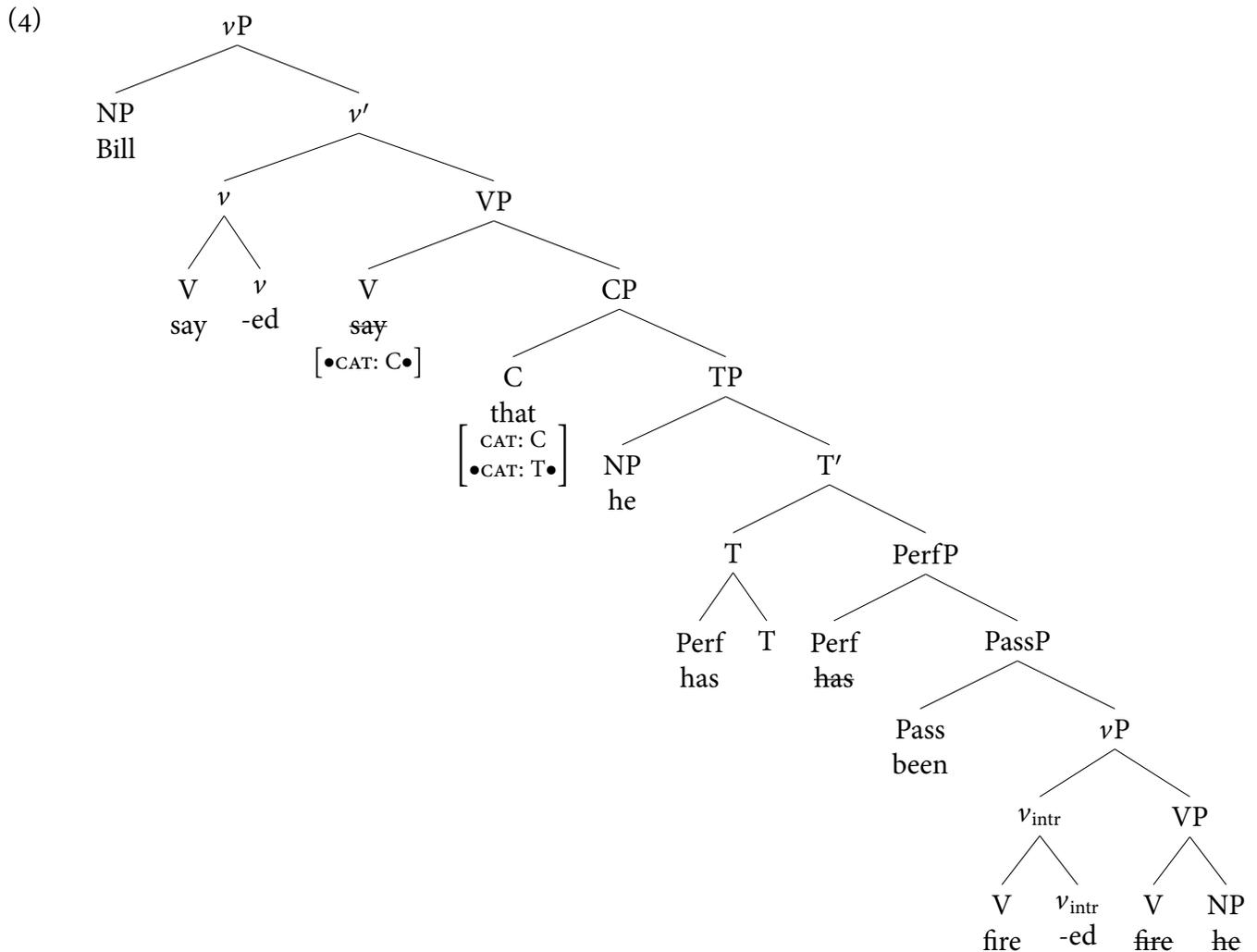
(2)



- Given our current assumptions, then sentences are TPs.
- The sentence above can be embedded under a verb like *say*. When it is, it may appear with an additional element *that* (3a).

(3) a. Bill said [that he has been fired]  
 b. Bill said [he has been fired]

- What category is this word? We will treat as belong to a distinct category, namely C standing for complementizer.
- The analysis of (3a) up to the higher  $vP$  is shown below:



- What about the sentence in (3b) which does not contain *that*? We have two options here: Either we could say that the entire CP projection is missing here and a verb like *say* may optionally select for a TP in addition to a CP. The alternative option, the one we will adopt here, is that the CP is structurally present, but its head is silent (much like is the case for T).
- This means we should add to the following two entries to our lexicon for complementizers:



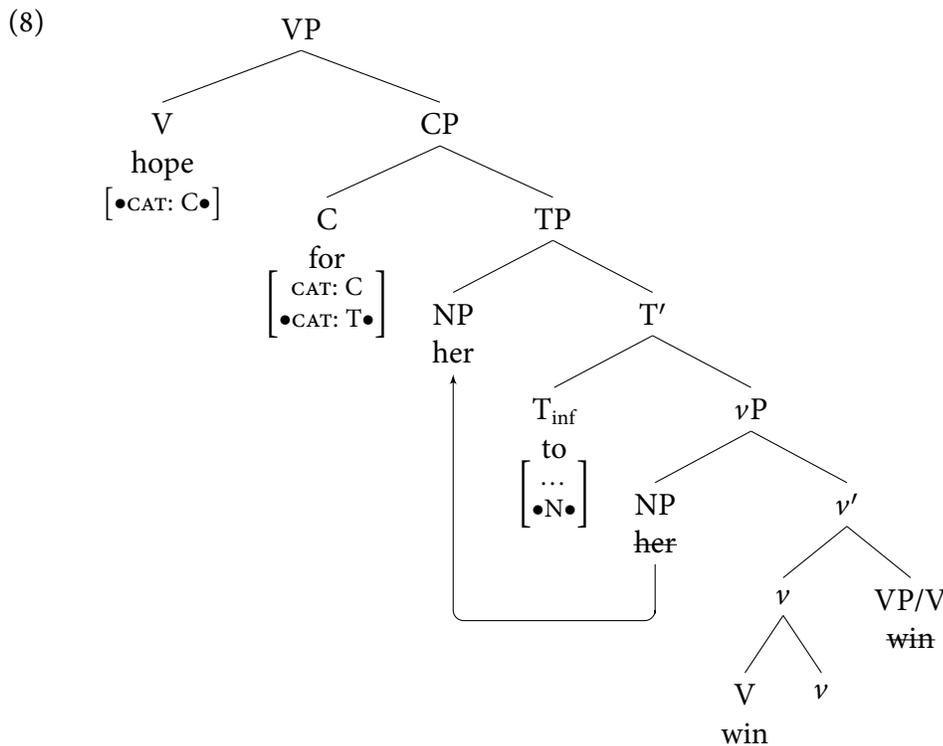
- Why should we assume that there is an empty C head? Well, there are multiple reasons. One is that we will see that this kind of empty C head actually has discernible functions in the clause. Second, we will also see that there are some cases in which CP is absent in an embedded clause, but that this is probably not true here.
- In order to see a possible function of null C, consider the following sentences:

- (6) a. I had hoped for her to win.
- b. I had hoped to win.

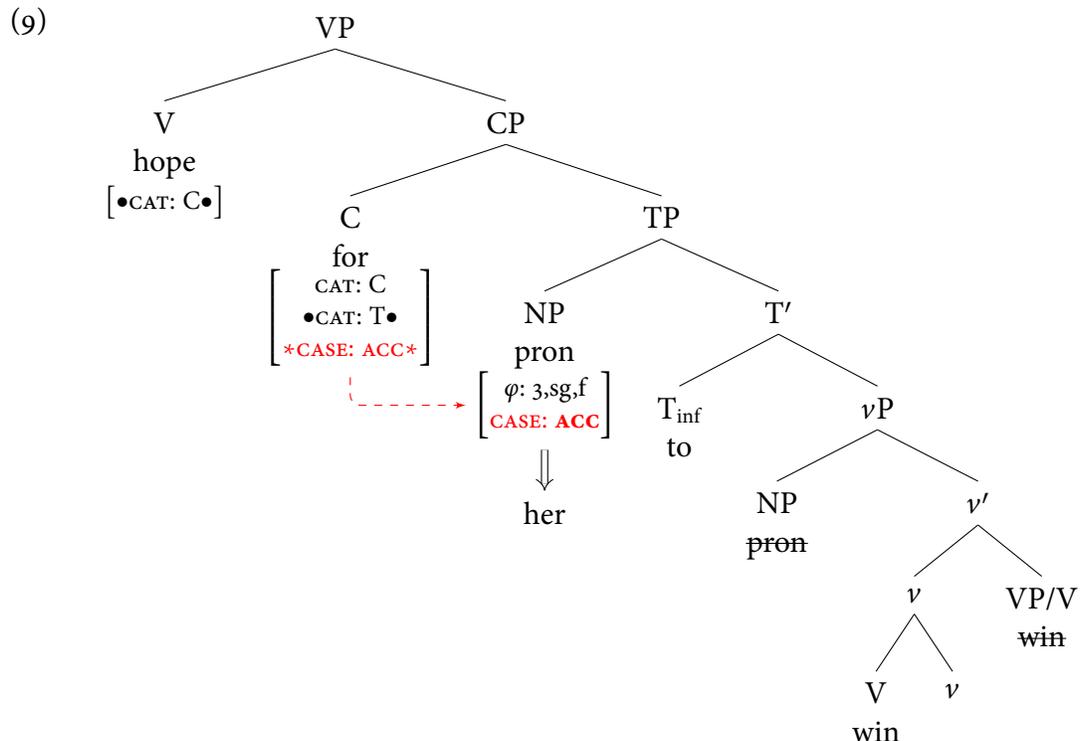
- In (6a), the subject of the verb *win* is overt. Interestingly, the 3rd singular feminine pronoun surfaces not in the normal case for subjects (nominative), but with the case we normally find on objects (accusative). Thus, we get *her* instead of *she*. This indicates that T is not assigning case in this kind of clause, as it normally would.
- Furthermore, this kind of embedded sentence shows no agreement with the subject. As we have already mentioned, we can assume that T is realized as *to*. Given this, we can assume that this T, the one that we find in non-finite or ‘infinitival’ clauses has a distinct lexical entry and category ( $T_{inf}$ ) as shown below:

$$(7) \left[ \begin{array}{c} T_{inf} \\ to \\ CAT: T_{inf} \\ \bullet CAT: \left\{ \begin{array}{l} v, v_{intr}, Perf, Pass, \\ Prog, Mod, Neg \end{array} \right\} \bullet \\ \bullet N \bullet \end{array} \right]$$

- This head selects the same range of complements as finite T, but does not assign case or tense or  $\varphi$ -features. It does, however, still have a  $[\bullet N \bullet]$  feature which will trigger movement of the subject to the specifier of T (or can be checked by merging *there* as we will see a bit later).
- With this in mind, the analysis of the embedded clause in *I had hoped for her to win* is shown below:



- Here, we assume that *for* occupies the same position in the clause as *that* above; it is a C head.
- This derives the correct order, but we still have to account for why the subject here shows up in accusative case.
- Perhaps, non-finite T could be assigning accusative (we will see why this probably isn't the case), but the alternative we will go with for now is that it actually the C head *for* bears an accusative case probe and can assign it to the subject of the embedded infinitival clause:



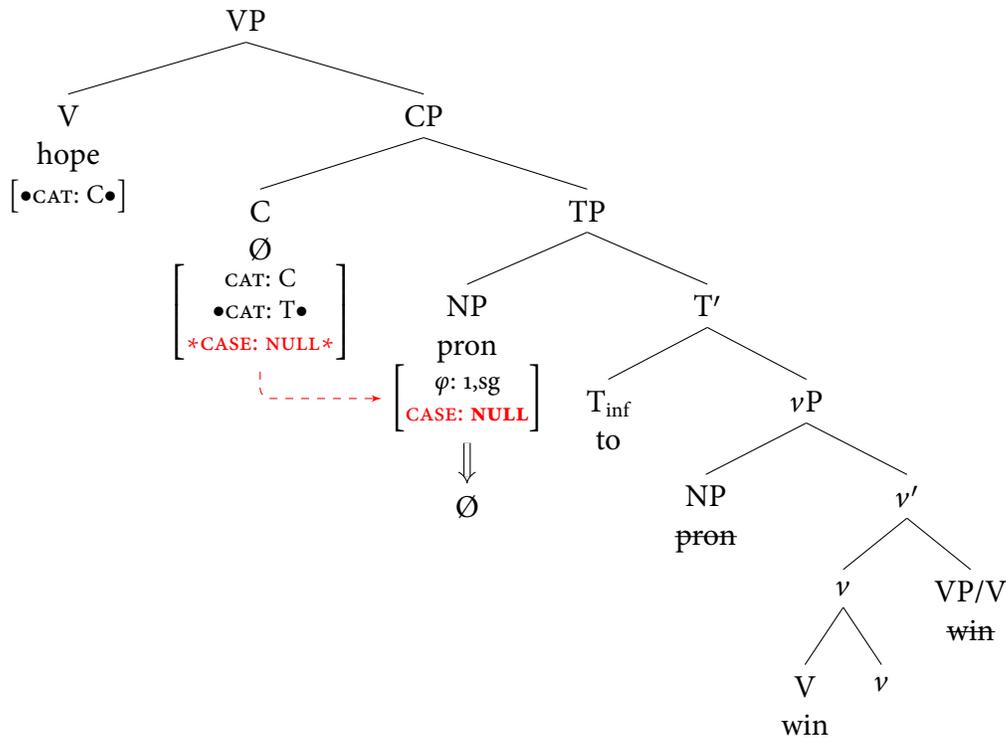
- Now let's turn our attention to the second example in (6), *I hope to win*. There are a couple of important differences to note here: the complementizer cannot be overt *for* (10a) and neither can the embedded subject (10b).

- (10) a. \*I had hoped for to win.  
b. \*I had hoped you to win.

- Nevertheless, there is an interpreted subject for the verb *win*. In fact it has to be same as the overt subject of the higher verb.
- So, how can we account for that fact there is a correlation between having an overt vs. null subject and having/lacking a complementizer? We can posit two distinct C heads.
- For the sister of *hope*, we can choose either the C head realized as *for*, in which case the subject will be an overt subject in accusative case, or we can choose a null C head in which case the subject will also be null.

- Why is the subject obligatorily null though? C does not select the subject of the embedded verb, so we cannot encode this relation through selection in a straightforward way (e.g. by saying that *v* must select some special null pronoun). Instead, we can view this as a kind of Case. Let's posit another value for case, namely [CASE: NULL], and assume that any phrase with this value will be realized as null/the empty string ( $\emptyset$ ) by our morphological rules, regardless of its other features.
- On this assumption, the special null C we have been talking about will assign 'null Case' to the subject as in (11).

(11)



- We therefore need to distinguish these two types of C that select for TP<sub>inf</sub>. The respective entries are shown in (12).



- Notice that I have assigned these different sub-categories. Why did I call the null version C<sub>ctrl</sub>, well this is due to another property that this C head has. Not only must the subject it assigns Case to be null, this subject must prefer to the subject of the higher clause.
- This is a phenomenon known as *control*. The empty subject is 'controlled', i.e. obligatorily co-referent with, an NP in the higher clause. In the case of *I hope to win*, it is the subject, so we call this *subject control*.

- Not all verbs allow for such structures, e.g. \*I believe to win, which is not possible with the meaning ‘I believe that I will win’ (compare I hope to win and I hope that I will win). We can therefore say that hope is a subject control verb since it permits such a structure. A more technical way of saying this is that it may select for C<sub>ctrl</sub>.

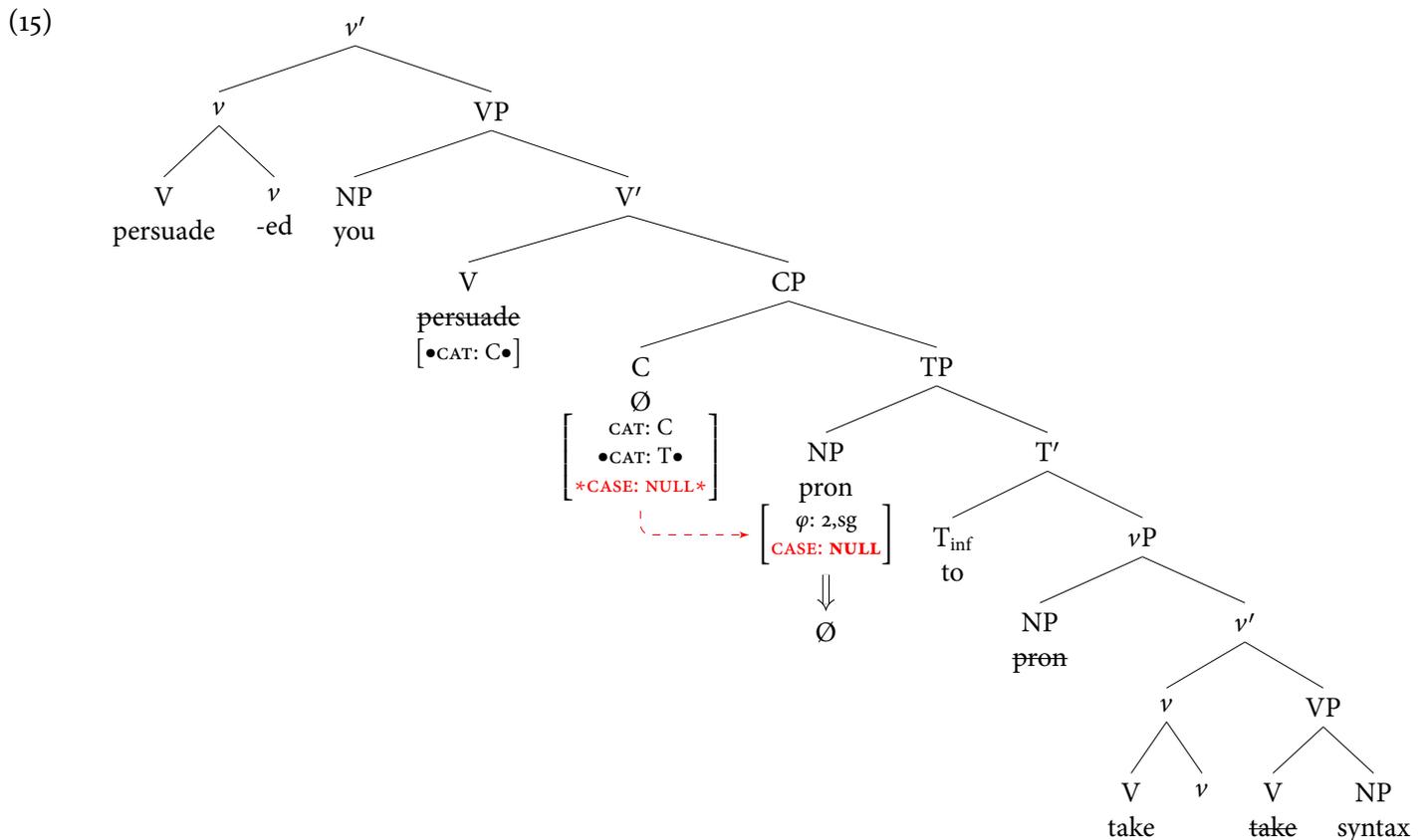
- It is common practice to represent the embedded subject in a control structure as PRO<sub>i</sub> where the subscript indicates co-reference:

(13) I<sub>i</sub> hope [PRO<sub>i</sub> to win]

- We will not try to encode this obligatory coreference property in our grammar. How and where exactly this should be encoded in the grammar is still a matter of some debate.
- So, now we have seen that there is subject control one might wonder if there is a comparable type of verb that requires PRO to refer to the object. As the following examples show, there are indeed so-called *object control verbs*:

(14) a. I persuaded you<sub>i</sub> [PRO<sub>i</sub> to take syntax]  
 b. I advised you<sub>i</sub> [PRO<sub>i</sub> to take syntax]

- We can assume that these have exactly the same syntactic structure, except for the fact that the higher verb selects an additional argument. The tree below shows this (omitting some details):



- We can now revisit a problem that I brought up in the very first set of lecture notes.
- Back then, I pointed out that changing the verb in the following two sentences radically alters the meaning. The pronoun *herself* must refer to *Sally* in (16a), but refers to *Jane* in (16b).

- (16) a. Jane convinced Sally to take herself more seriously.  
 b. Jane promised Sally to take herself more seriously.

- It seemed like quite a puzzle why changing the verb should have this effect. In fact, recall what our rule about *self*-pronouns was: They have to refer to closest matching c-commanding NP. This condition is violated in (16b), since the closest matching referent is *Sally*. So, this is also looks like a problem for that generalization on the face of things.
- But given what we have said about some verbs being control verbs, the data above actually make perfect sense once we take into account that *convince* is an object control verb and *promise* is a subject control verb. In that case, we will actually have the structural analyses below:

- (17) a. Jane convinced Sally<sub>1</sub> [PRO<sub>1</sub> to take herself<sub>1</sub> more seriously]  
 b. Jane<sub>2</sub> promised Sally [PRO<sub>2</sub> to take herself<sub>2</sub> more seriously]

- Now, the closest matching c-commanding antecedent for *herself* is actually the null subject of the embedded clause PRO. What this PRO refers to is determined by the ‘control’ specification of the verb. This is why the reference of *herself* is sensitive to the type of verb we have because it in turn changes the reference of the silent subject we have been assuming.
- For this reason, adopting an analysis of this kind where the subject of non-finite clauses is silent, yet syntactically represented allows us to understand the effect we observed in (17).
- So to take stock, we have now seen three types of non-finite embedded clauses: There are those with *for* and an overt subject (18a) and those with null subjects that either refer to the higher subject (subject control) (18a) or the higher object (object control) (18c).

- (18) a. I had hoped [for you to win]  
 b. I<sub>i</sub> had hoped [PRO<sub>i</sub> to win]  
 c. I had asked you<sub>2</sub> [PRO<sub>2</sub> to win]

- So, we can divide up the clause-embedding verbs of English into these three classes. Inevitably, some verbs will be able to show up in more than one class (just like some verbs can be both transitive and intransitive verbs). We have already seen that *hope* can take both a *for*-infinitive complement and a subject control complement.
- These heads will have different selecting properties, e.g. *hope* will have [ $\bullet\{C_{for}, C_{emb}, C_{ctrl}\}\bullet$ ], meaning that it can select either a *for to*-infinitive, a finite embedded clause (with or without *that*), or a non-finite (subject) control clause. Other verbs are more restricted in their selecting potential.

- **NOTE: The final part of this section is additional information that we did not cover in class. You are not required to know it for the homework or final, so only read on if you are actually interested. Section 2 is very relevant for the homework/final though ☺.**

- For a verb like *want*, we might be tempted to characterize it as an optional subject and object control verb (unlike *hope* for example):

- (19) a. I want to leave.  
b. I want you to leave.

- Furthermore, verbs like *believe* would just be an object control, but not a subject control verb:

- (20) a. \*I believe to be the best.  
b. I believe her to be the best.

- It turns out that things are a little more complicated, however. Despite first appearances, these verbs are not object control verbs. What appears to be the object of *want* and *believe* is actually the subject of the embedded clause in the specifier of TP:

- (21) a. I want [you to leave]  
b. I believe [her to be the best]

- I won't dwell on this point too much, but here is one reason we can tell they are different.
- Remember that we can insert *there* into the specifier of T in order to check its [ $\bullet$ N $\bullet$ ] feature. But *there* can't just be inserted in place of an argument, e.g. a subject (22b).

- (22) a. There were three people being arrested outside.  
b. \*There were/was arresting three people.

- For this reason, we can explain the following difference:

- (23) a. I wanted/expected/believed [there to be a party tonight]  
b. \*I asked/persuaded/convinced there<sub>i</sub> [PRO<sub>i</sub> to be a party tonight]

- Since the verbs in (23b) are genuine object control verbs, *there* would have to be the object of the verb, which is not possible for various reasons (it also doesn't refer to anything so it cannot 'control' PRO). It is possible for the verbs in (23a) precisely because they are not object control verbs, they are verbs that allow an overt subject in their specifier (this is also true of *for*-infinitives to: *I had hoped for there to be a party tonight*).
- There are some details that we won't go into here, but the basic idea is that the subject of the embedded clause is getting its case from the higher  $\nu$  directly: as such they are called *exceptional case marking* or *ECM* verbs. Feel free to read up on this distinction if you are interested. For present purposes, we won't have too much more to say about it.

## 2 Questions

- Let's recap what we have seen so far. We have now posited four distinct C heads for embedded clauses, repeated below, as well as a new T head for non-finite clauses.

(24) a.	$C_{emb}$ that $\left[ \begin{array}{l} \text{CAT: } C_{emb} \\ \bullet \text{CAT: T} \bullet \end{array} \right]$	c.	$C_{emb}$ $\emptyset$ $\left[ \begin{array}{l} \text{CAT: } C_{emb} \\ \bullet \text{CAT: T} \bullet \end{array} \right]$
b.	$C_{for}$ for $\left[ \begin{array}{l} \text{CAT: } C_{for} \\ \bullet \text{CAT: T}_{inf} \bullet \\ * \text{CASE: ACC} * \end{array} \right]$	d.	$C_{ctrl}$ $\emptyset$ $\left[ \begin{array}{l} \text{CAT: } C_{ctrl} \\ \bullet \text{CAT: T}_{inf} \bullet \\ * \text{CASE: NULL} * \end{array} \right]$

- For now, let's just focus on embedded finite clauses, i.e. the heads in (24a) and (24c). We have seen that both of these heads are possible in an embedded clause. But what about a main clause?
- We know that (24a) cannot occur in non-embedded sentences:

(25) \*That he has been fired

- But what about the null variant might be useful in accounting for something we have already talked about this quarter: the formation of so-called 'yes-no questions' (or 'polar questions') such as (26b).

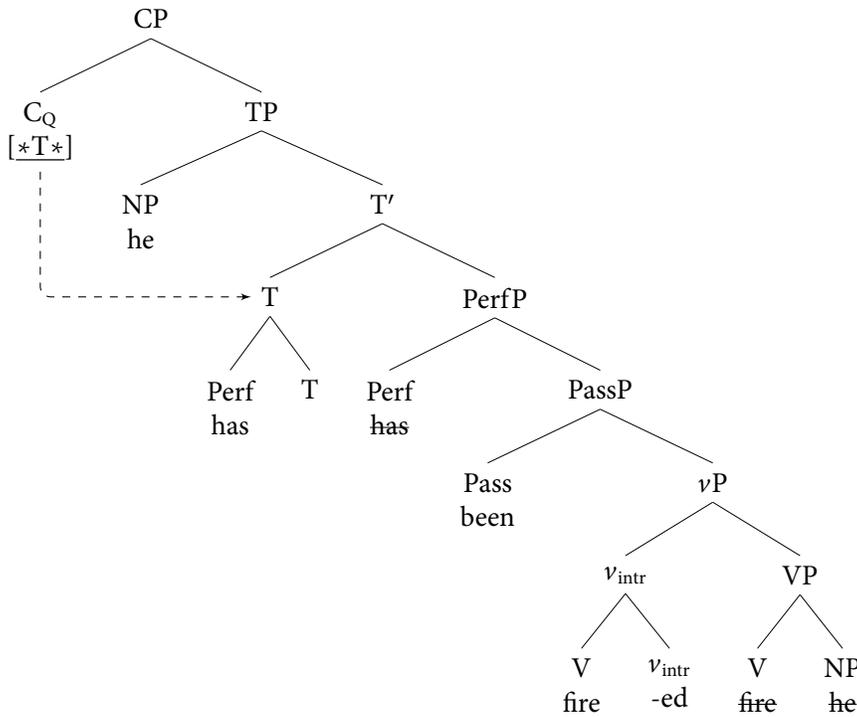
(26) a. He has been fired.  
b. Has he been fired?

- The word order in (26b) is typically referred to as *subject/auxiliary inversion*. To derive it, we need some kind of position above the subject for the highest auxiliary to move to.
- If we have a CP in higher clauses (also called 'matrix clauses'), then we can assume that the highest auxiliary moves there:

(27) a. Has he \_\_\_ been fired?  
  
 b. Hasn't he \_\_\_ been fired?  
  
 c. Will they \_\_\_ fire him?  

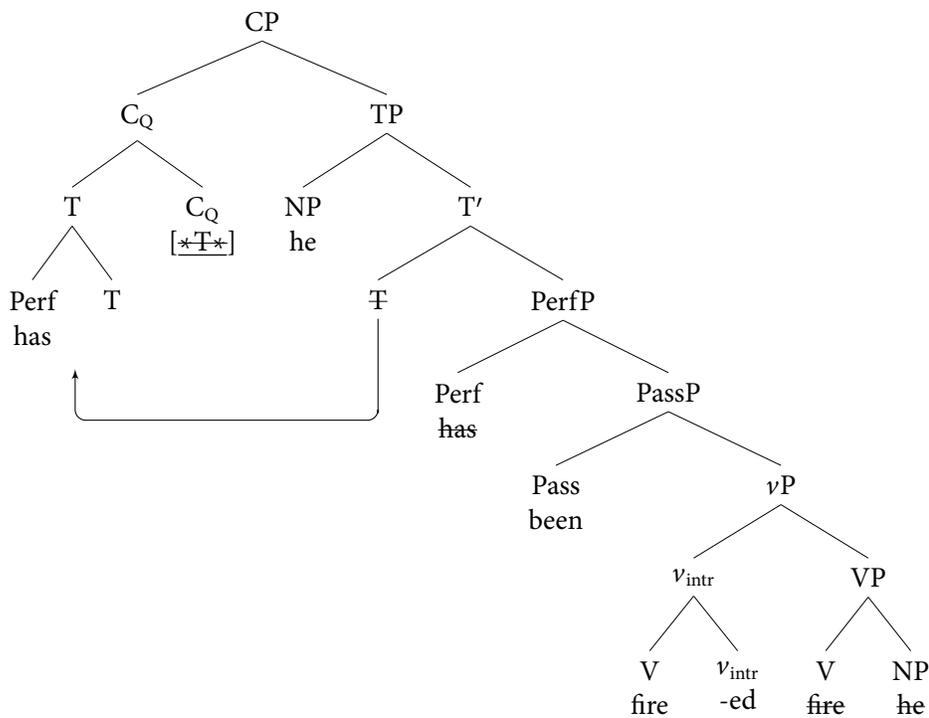

- What features do we need? This would suggest that there is a strong feature [ $*T*$ ] on the C head in such questions:

(28)



- As a result of Agree for this feature, T (and anything adjoined to T) will move to C:

(29)



- Is there a C in non-questions? We could definitely assume this so that all sentences are actually CPs rather than TPs. This C would have to be different though, it would just select a TP and not do anything else, e.g.

triggering head movement (though there may well be other languages where this is the case, hint hint).

- So, we have now expanded our inventory of C heads even further:

(30)	a.	$  \begin{array}{c}  C \\  \emptyset \\  \left[ \begin{array}{l} \text{CAT: } C \\ \bullet \text{CAT: } T \bullet \end{array} \right] \\  C_{\text{emb}} \\  \text{that} \\  \left[ \begin{array}{l} \text{CAT: } C_{\text{emb}} \\ \bullet \text{CAT: } T \bullet \end{array} \right] \\  C_{\text{for}} \\  \text{for} \\  \left[ \begin{array}{l} \text{CAT: } C_{\text{for}} \\ \bullet \text{CAT: } T_{\text{inf}} \bullet \\ * \text{CASE: ACC} * \end{array} \right]  \end{array}  $	d.	$  \begin{array}{c}  C_Q \\  \emptyset \\  \left[ \begin{array}{l} \text{CAT: } C_Q \\ \bullet \text{CAT: } T \bullet \\ * \text{CAT: } T * \end{array} \right] \\  C_{\text{emb}} \\  \emptyset \\  \left[ \begin{array}{l} \text{CAT: } C_{\text{emb}} \\ \bullet \text{CAT: } T \bullet \end{array} \right] \\  C_{\text{ctrl}} \\  \emptyset \\  \left[ \begin{array}{l} \text{CAT: } C_{\text{ctrl}} \\ \bullet \text{CAT: } T_{\text{inf}} \bullet \\ * \text{CASE: NULL} * \end{array} \right]  \end{array}  $
	b.		e.	
	c.		f.	

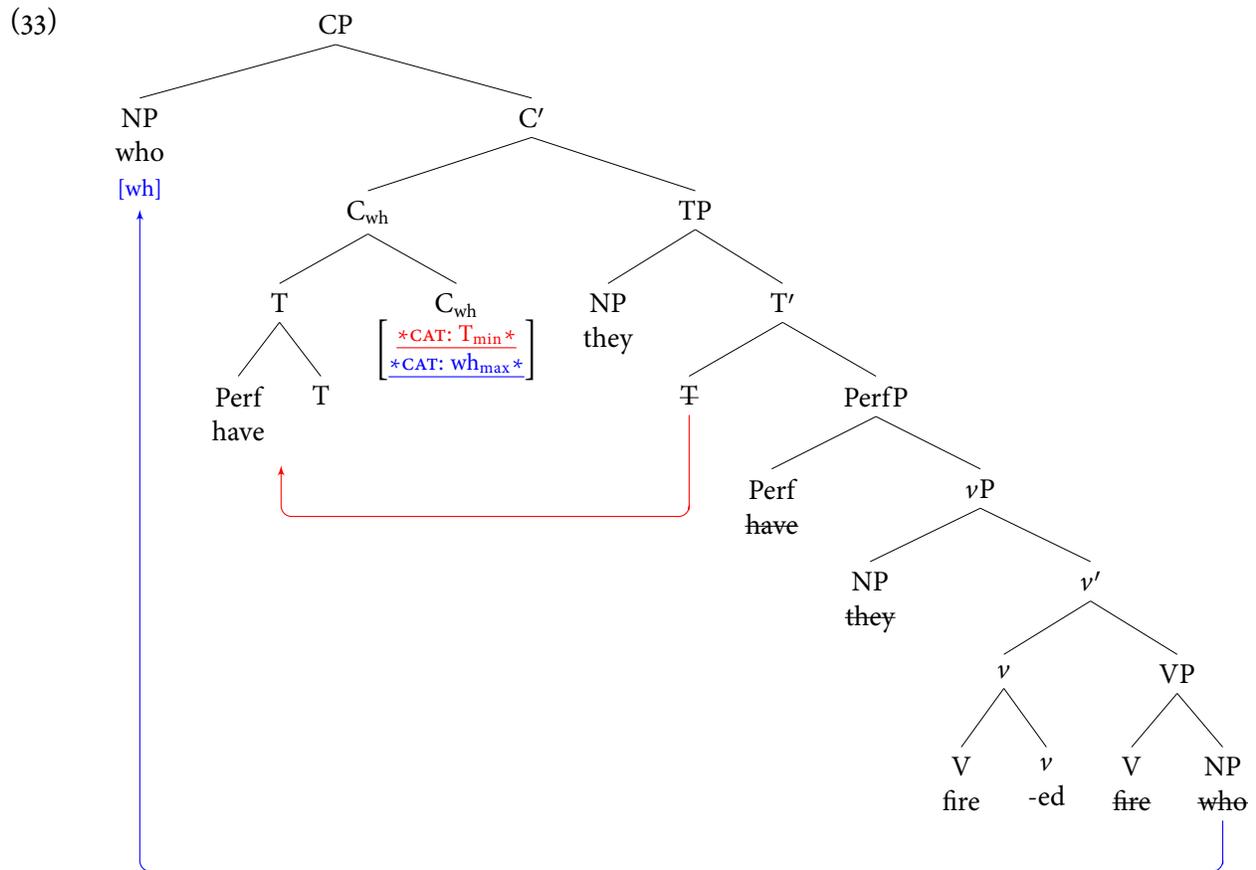
- Are there other cases in which we find subject/auxiliary inversion? Yes, we find it with a different kind of question that involve a question word like *who* or *what*. These are called *wh-questions*.

- (31)
- a. They have fired Bill
  - b. Who have they fire?
  - c. \*Who they have fired?

- A particular property of this kind of question is that the have a question word *who* has to be at the beginning of the sentence:

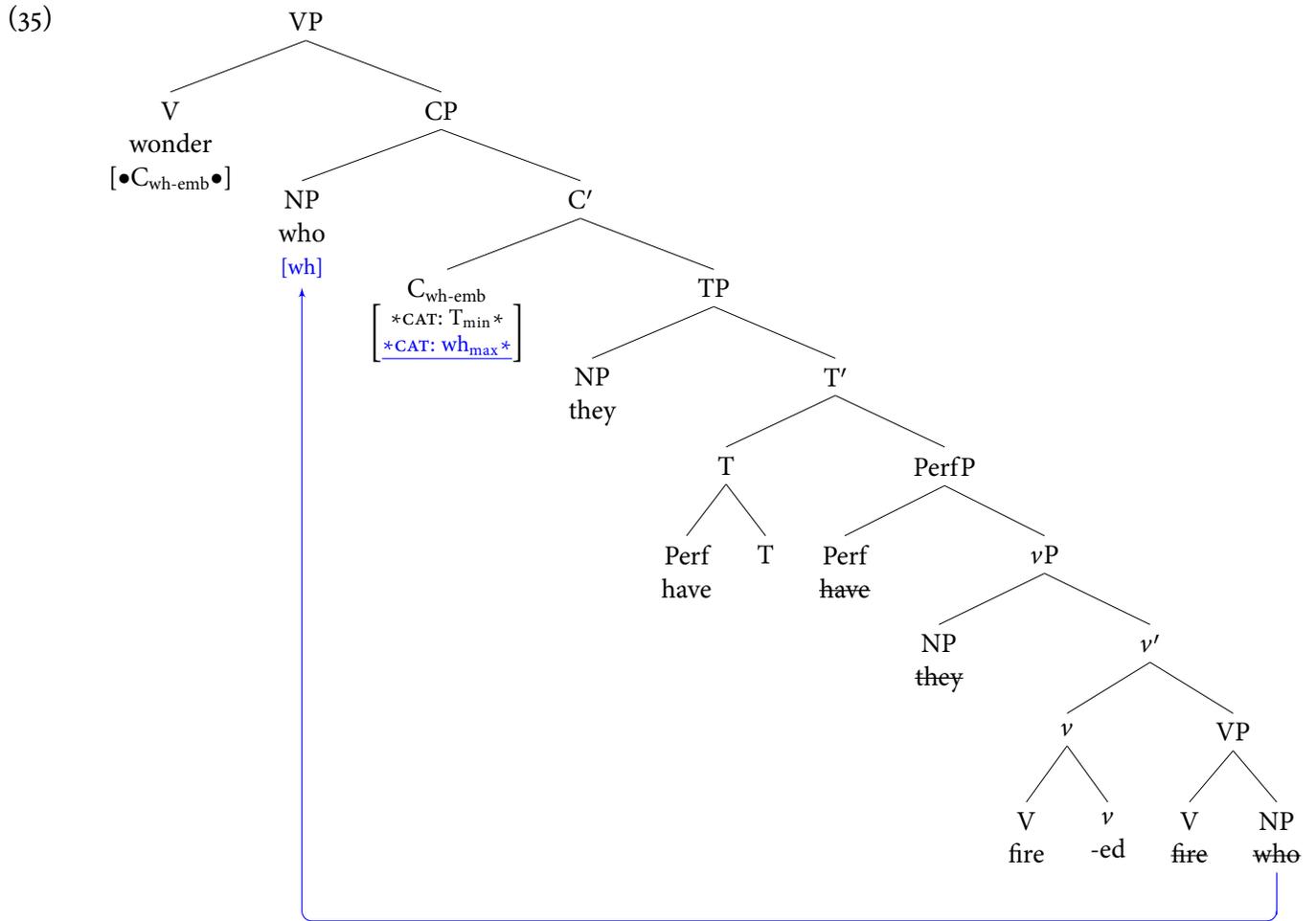
- (32) \*They have fired who?

- We need to encode this fact somehow. One possibility is to say that C has to select an NP, too. But this won't be enough, since what moves needs not just be a noun phrase, but also a particular kind (a 'question word'). Instead, we can assume that wh-words have a feature [wh] that single them out as such. We can add the corresponding probe feature [\*wh\*] to C in wh-questions.
- On its own, however, this will just make sure that the sentence contains a wh-word, as in (32). We also need to ensure that there a wh-word in the sentence **and** that is moves to the specifier of C. We could make this feature 'strong' in the same way we did for features
- A piece of notation that could do this would be: [\*F<sub>min</sub>\*] for features that trigger movement of a minimal projection (a head) and [\*F<sub>max</sub>\*] for features that trigger movement of a maximal projection (a phrase).
- On this view, the C head we have in questions with the feature [\*wh<sub>max</sub>\*] to trigger movement of some phrase to the specifier of C in addition to [\*T<sub>min</sub>\*] to trigger subject/auxiliary inversion. If we put these things together, we get the following:

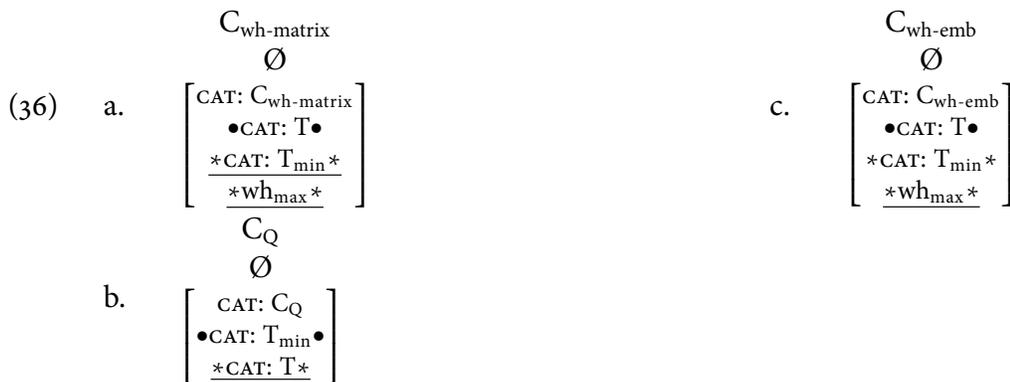


- The C head in this tree also *requires* that the structure contains a *wh*-word. If it didn't, then the structure would be ungrammatical. We cannot use this head to derive *Have they fired him?* then. This is why I have labelled it 'C<sub>wh</sub>'.
- It is also possible to have *wh*-questions as embedded clauses:
 

(34) a. I wonder who they have fired.  
 b. \*I wonder who have they fired.
- Importantly, subject-auxiliary inversion here is ungrammatical (34b). Nevertheless, there is still movement of a *wh*-word in embedded *wh*-questions. How can we account for this difference? Well, the observation that there is no subject-auxiliary inversion in embedded clauses translates into a lack of T-to-C movement in our theory. So, this means that the [*\*T\**] feature on C must be weak in embedded clauses:



- As a result, T does not move to C and therefore does not carry the auxiliary adjoined to T up to the C head.
- It seems that we need to distinguish between two kinds of C heads that derive wh-questions.
- The three heads we have assumed for wh-questions in English are given below:



- The distinction is not only important for capturing the distinct properties they have, but also making

sure that verbs select the right kind of structure. A verb like *wonder* may only select  $CP_{wh-emb}$ , but not  $CP_{wh-matrix}$ . If the latter were possible, we would get the ungrammatical sentence (34b).

- As ever, all of this effort and technical machinery has a clear purpose: We want to derive the structures we find and rule out the structures we don't.