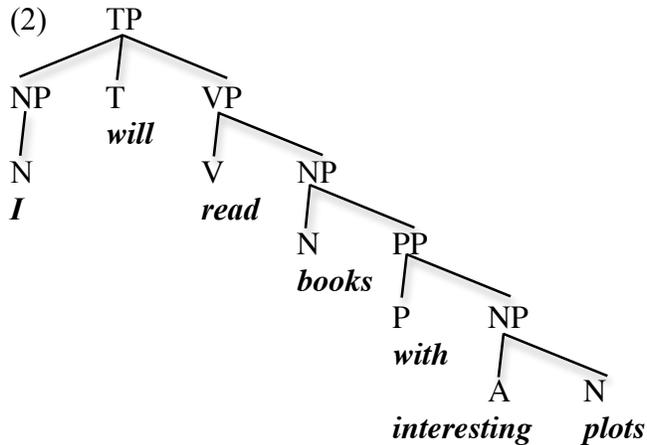


## 24.902, Handout 1: More about Phrase Structure

Last time we'd developed something like the following set of phrase structure rules:

- (1)  
TP → NP T VP  
VP → V (NP) (PP) (CP)  
NP → (Det) (A) N (PP)  
PP → P (NP)  
CP → C TP

Put these together with a lexicon (which lists words that count as instances of N, A, P, and so on), and you can generate trees like:



The little mini-grammar in (1) has a couple of desirable properties already: it has recursion, and it can handle certain cases of ambiguity [remember?]

And yet, the mini-grammar in (1) is not the final word in grammar. It has at least three serious problems:

- (3)
- they're quite unrestrictive; nothing prevents us from writing a rule like VP → NP TP, if we want
  - they're largely redundant with things we're going to have to state in the lexicon anyway (we talked about this some, and we'll come back to it)
  - they turn out not to make enough structural distinctions (more about this later, too)

Consider the VP rule:

VP → V (NP) (PP) (CP)

We said before that we're going to combine these rules with lexical entries to get fragments of English. Lots of lexical choices won't work out, of course:

- (4) a. The dragon **devoured** the villagers  
b. \*The dragon **devoured**
- (5) a. The dragon **rejoiced**  
b. \*The dragon **rejoiced** his victory
- (6) a. The dragon **put** the peasant upon the plate  
b. \*The dragon **put**  
c. \*The dragon **put** the peasant  
d. \*The dragon **put** upon the plate

The data in (4-5) are described by a classic observation: some verbs are *transitive* (they require objects) and others are *intransitive* (they can't have objects). And, in fact, as we see in (6), there are verbs with even more extensive requirements.

So in the lexical entry for *put* we'll need to have the information "needs an NP and a PP". Once we have that information there, do we really want to restate it in a phrase structure rule?

--> **selection**, which we'll be talking about more later. Heads get to **select** (or **subcategorize**) for certain types of phrases.

Last time I floated the idea that we could replace all of our phrase structure rules with something like "XP → X plus whatever the lexicon says X needs". But that turns out not to quite work, because not everything is selected:

- (7) a. The dragon devoured the villagers (**yesterday**) (**in Omaha**)  
b. The dragon rejoiced (**yesterday**) (**in Omaha**)  
c. The dragon put the peasant upon the plate (**yesterday**) (**in Omaha**)

phrases like *the villagers*, *the peasant*, and *upon the plate*, which are picky about which heads they cooccur with, are **complements** of the heads; we'll say that the heads **select** (or **subcategorize**) for them.

phrases like *yesterday* or *in Omaha*, which aren't selected, are **adjuncts**.

Now for problem (3c); these phrase structure rules give us an insufficiently fine-grained structure: [judgments and \*'s removed for audience participation purposes]

- (8) a. John decided on the boat  
b. John decided on the boat in Omaha  
c. John decided in Omaha on the boat  
d. John decided on the boat on the plane

-->complement/adjunct distinction has structural consequences

(9) John decided on the boat, and Mary did so on the plane

- (10) a. a student of physics from Brazil  
b. a student from Brazil of physics

- (11) a. a Brazilian physics student  
b. a physics Brazilian student

- (12) a. ...a student from Rwanda, and one from Brazil  
b. ...a student of chemistry, and one of physics

- (13) a. ...a Rwandan student, and a Brazilian one  
b. ...a chemistry student, and a physics one

- (14) a. an alleged Brazilian student  
b. a Brazilian alleged student

- (15) a. an alleged student from Brazil  
b. alleged student of physics

- (16) a. John knocked on the door twice three times.  
b. John knocked on the door three times twice.

tests; how do we distinguish between complements and adjuncts?

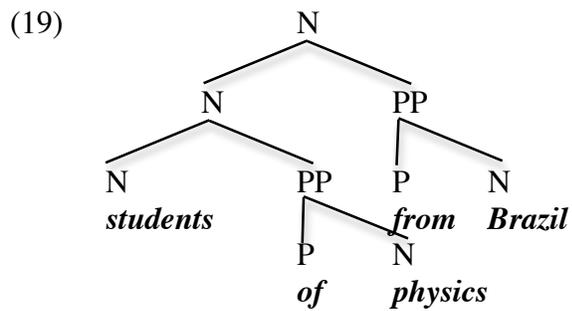
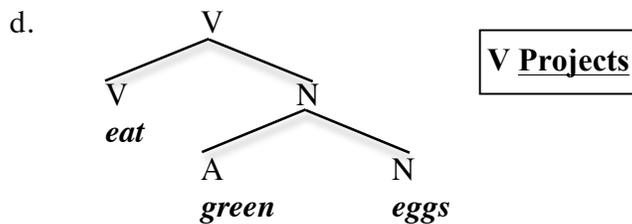
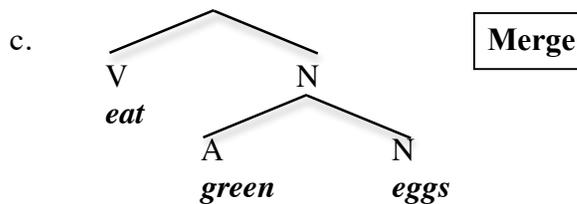
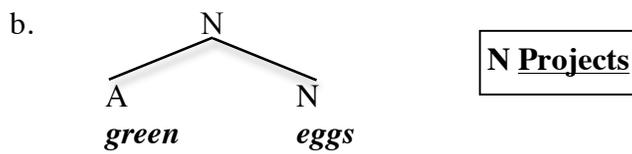
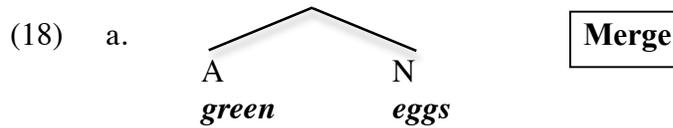
- obligatoriness?
- relative scope (ambiguity with *alleged*)
- ordering (student with green hair of physics, student of physics with green hair)
- *one*-replacement, *do so*-replacement
- pickiness (student/\*orange of physics)
- particular choices of P? (a student of impeccable character)
- extraction:

- (17) a. What subject is she a student of?  
b. What color hair is she a student with?

**Tree construction algorithm (Merge):**

- (i) Take two things, A and B, and put them together to form a third thing, C.
- (ii) Give C the label of either A or B
- (iii) Repeat as necessary.

[crucially, A and B may be either lexical items, or previous products of Merge. That is, Merge reapplies to its own output. In other words, Merge is **recursive**]



These trees resemble the trees generated by the **X-bar schema**: compare and contrast...

- **XP=maximal projection** (node with label X which is not dominated by any projections with labels X from the same head (terminal node))
- **X=head** (node which dominates no other nodes)
- **X' ("X-bar")=intermediate projection** (node which is neither a head nor an XP)

a new constituency test; **coordination**

- (20) a. John bought [an apple] and [a lemon]  
b. Mary [bought an apple] and [sold a lemon]

(handle with care: **Right Node Raising**)

- (21) a. [John bought], and [Mary sold], a shiny new lemon  
b. [John bought], and [Mary sold], it

- (22) I [have won the lottery] and [will never work again]

Now, our discussion has mostly been about English. Does every language need all this structure?