September 15 & 17 class summary: 24.902

The week actually began with some unfinished business from last week, which is contained in the previous summary. Then we moved on...

1. Preliminaries

A word often cares about its syntactic environment. A verb, for example, may require two, one or zero complements, and (seemingly -- see below!) may specify the syntactic category of its complement.

Examples: *put* requires NP PP, *devour* requiresNP, *depend* requires PP, *eat* takes an optional NP.

(1) a. Sue put the book under the table.
   b. *Sue put the book.
   c. *Sue put under the table.

(2) a. John devoured the sandwich.
   b. *John devoured.
   c. *John devoured under the table.

(3) a. Bill depended on his tablet PC.
   b. *Bill depended from his tablet PC.
   c. *Bill depended his tablet PC.
   d. *Bill depended.

(4) a. Sue ate the sandwich.
   b. Sue ate.

This information forms part of the lexical entry for a verb, and is partly arbitrary.

2. Chomsky's subcategorization mechanism and his constraint

Chomsky's (1965) mechanism*:*
   How words get inserted into structure:
   a. Step 1: Nodes such as V, N, A, etc. undergo a *subcategorization rule* that assigns them to subcategories, whose name describes the environment in which the rule applies. The feature on the right of the arrow, by the way. was called a "Complex Symbol" (CS):

   Example: V--> [+ __ NP PP] / __ NP PP

   b. Step 2: A CS is replaced by a lexical item of the appropriate category that also bears the same subcategorization feature. For example, "[+ __ NP PP]" might be replaced by *put*, because *put* bears the feature [+ __ NP PP]. It

could not be replaced by *depend*, because *depend* bears the distinct feature [+ __ NP PP].

OK, painful detail. What is interesting to us is a constraint proposed by Chomsky on the ability of a subcategorization rule to control the environment in which it applies.

Chomsky claimed that the environment of the subcategorization rule for X makes reference to *all and only the sisters of X*. More simply still: the subcategorization property of X cares only about the sisters of X.

Terminological note: As mentioned in class, a slightly different way of speaking has arisen in the field of syntax. We say that a verb like *put* "subcategorizes for" an NP and a PP. Likewise a verb like *eat* "subcategorizes for" an optional NP.

Chomsky's constraint has interesting implications for *language acquisition*. Granted that a child must maintain some record of the syntactic environment in which lexical items occur (or else subcategorization information would not be acquired), Chomsky's constraint suggests that this information is quite limited. The child does not attend to the entire tree when keeping these mental records.

That said, it is not clear why subcategorization phenomena exist in the first place. But they do.

An interesting point that was came up (inadvertently) in class: the child must already know a lot about syntactic structure to knowe what information to ignore when creating lexical entries!

3. Chomsky's mistake

Chomsky's constraint is empirically on the right track: we do not find obligatory modifiers or other adjuncts. Where a constituent is obligatory, it appears to be a complement. For example:

(5)   *Sue depended on her friends and Bill did so on his relatives.

But it is also partly wrong. Let us focus on PP complements.   Contrary to what Chomsky's (1965) proposal leads one to expect, many verbs actually ask for particular prepositions in their PP complements -- not just any old PP. Thus, *depend* requires *on* in English, not *from*:

(6)   a. Bill depended on his friends.
      b. *Bill depended from his friends.

Similarly, the noun *preference* (we didn't do this one in class), when it takes a PP complement (which it doesn't have to do) requires that the preposition of the PP be *for*:

(7) a. Mary's preference for opera
   b. *Mary's preference of opera / to opera / etc.

Likewise, the adjective *fond* requires a PP complement, whose preposition must be *of*:

(8) a. John is fond of his children.
   b. *John is fond for his children / at his children / etc,
   c. *John is fond. (ok in older English, with different meaning)

Predictability: The semantic class of preposition required by a particular verb, noun or adjective restricts the actual choice of preposition somewhat in these cases. It might be surprising to find *from* selected by *fond*, for example. But the actual preposition, or selection for a preposition in the first place, is unpredictable. The comparison between English and Russian makes this clear. Note that *from* is actually the preposition used in Russian with 'depend':

(9)   English   Russian
   depend on   [*zaviset' ot* 'from']
   speak to   [*govorit' s* 'with']
   looked at the wall [*smotret' na* 'on']
   look for the book [*iskat' knigu* - no preposition]

So we must modify Chomsky's proposal, using a term that picks out the P of a PP. The term we use is **head**. The P is the *head* of PP.

(10) **Chomsky's Constraint, revised (informal version)**
   The subcategorization feature of a lexical item α dictates the *heads* of the complements of α.

We will dispense with the extra mechanisms of Chomsky's approach, and just assume that nodes such as V bear the relevant subcategorization features, with *lexical insertion* below such nodes as V limited to verbs that match V in subcategorization features.

An additional observation (**not discussed in class**): It is worth noting that there is actually *no* verb (or noun or adjective) that requires a PP complement and doesn't care about the head of the PP. Verbs either name a particular head that a PP complement must have or else require a particular semantic class of PPs. *Put*, for example, requires both an NP and a PP complement, and demands that the head of the PP complement be a preposition that names the goal (endpoint) of the motion of the NP:

(11) a. Sue put the book on the table.
         [endpoint: surface of the table]
      b. Sue put the book inside the room.
         [endpoint: interior of the room]
      c. *Sue put the book from the bookcase.
         [*from* doesn't give the endpoint]

4. Using the constraint to identify heads

We can now use this generalization (assuming it is correct) to probe for heads of phrases.

C and CP

Consider, for example:

(12)   a. Mary believes [that Sue can read the book].
       b. Mary arranged [for Sue to read the book].

(13)   a. *Mary believes [for Sue to read the book].
       b. *Mary arranged [that Sue can read the book].

The verb *believe* allows a clause as a complement when the clause is introduced by the complementizer *that*. The verb *arrange* allows a clause introduced by the complementizer *for*. But *believe* is incompatible with *for* and *arrange* is incompatible with *that*. Looks like subcategorization!

This suggests that C (complementizer) is the head of an embedded clause. Following our terminological conventions, we will call embedded clauses CP.

An important note on complementizer *for*: It is crucial to distinguish complementizer *for* from preposition *for*. The complementizer *for* is followed by a sentence. An example of complementizer *for* is seen in (14). What is important here is the fact that the NP that follows *for* -- the pronoun *it* -- is the special near-meaningless *it* that can be used as the subject of weather verbs like *rain*:[1]
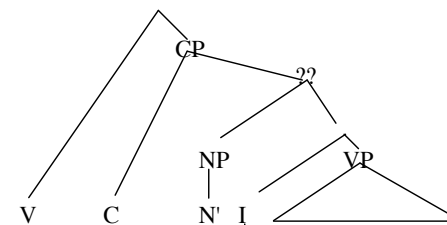
(14)   The student arranged [for   it   to rain on Thursday.]

When *for* is a complementizer, the sequence *for*+NP is not a constuent. One sign of this is the fact that it cannot be moved to the front of the sentence when focused:

---

[1] Called *expletive*, *pleonastic* or *dummy it*. All three terms are in use, and mean the same thing.

(15)   *For it, the student arranged to rain on Thursday.[2]

In contrast, when *for* is a preposition, the sequence *for*+NP may be moved:

(16)   a. We sang the song for Bill.
        b. For Bill, we sang this song.

And, of course, the preposition *for* may not take the special meaningless *it* as its object. If I come into the room say "it's raining", you won't ask me: "What's raining?". If I say (17) out of the blue, you *will* want to know what I was referring to with the word *it*!

(17)   We sang the song for it.

It's not unusual to find a preposition also doing duty as a complementizer. It's quite common for meaningful words to be "demoted" over time to function words (complementizers, determiners, etc.). The process is called *grammaticalization*, and there's a large (not always insightful, alas) literature on the topic.

I and IP

The same reasoning that led us to identify C as the head of the constituents *that Sue can read the book* and *for Sue to read the book* tells us that *to* is the head of the sister to C. The syntactic category of *to* is variously labeled as T (for Tense) or I (for Inflection). (Perhaps we'll take a vote in class on T vs. I. The reading by Carnie calls it T.)

The key observation is the fact that the complementizer *for* requires *to* (also known as an *infinitival* or *non-finite* marker), while the complementizer *that* requires anything but *to*.

(18)   **C cares about I**
        a. ...for Sue to read the book.
        b. *...that Sue to read the book
        c. *...for Sue can read the book / *...for Sue has read the book / etc.
        d. ...that Sue can read the book / *...that Sue has read the book / etc.

This tells us, probably, that the head of the mysterious "??" constituent in (14) is I, since C seems to subcategorize for it. The mysterious phrase itself — The sister to C — therefore, is called IP:

---

[2] Not actually starred if we allow a loony meaning in which we not only know something to which *it* refers, but students can meaningfully be said to rain.

(19)   The government arranged [for   it to rain on Thursday.]
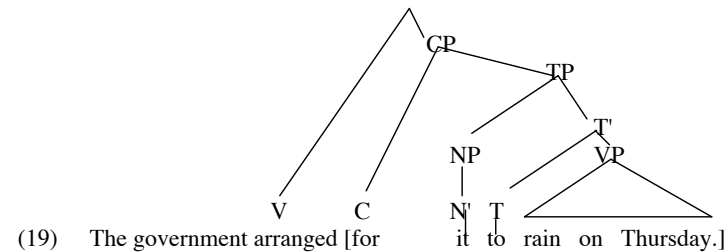
5. Japanese phrase structure

Consider now some sentences in Japanese whose word order is perceived by native speakers as "normal" -- or, in linguists' jargon, *unmarked*. [You'll see next week why I make this cautionary remark.] These sentences conform to a very neat generalization across all categories:

(20)   **Japanese word order**
        For any head H:
        a. H follows its complements.
        b. H' follows all adjuncts (modifiers).
        c. H' follows its specifier.

Expressing this using Phrase Structure rules, where *complements* stands in for whatever the complements of a particular head may be, etc:

(21)   **Phrase structure rules for Japanese**
        H' --> *complements* H
        H' --> *adjuncts*   H'
        HP --> *specifier*   H'

It is important (and reassuring) to note that the same elements singled out as heads when we examined subcategorization in English are identified as heads in Japanese according to the generalization in (20) and (21).

   Thus, for example, the fact that verbs in English whose complement is a a subordinate clause "care about" the particular choice of C told us that C is the head of the subordinate clause. Lo and behold, the C of a subordinate clause comes last in Japanese -- just where the head should be. That's why we call it CP!

   Similarly, the fact that English C "cares about" its complement's I, told us that I is the head of the complement of C. Lo and behold, the I of the complement to C comes last in Japanese. So we call the phrase IP.

The Japanese handout has many relevant examples. Here's the example that shows C final in CP and I final in IP (as well as V final in VP etc.):

(22)

```
                                    IP
                                   /  \
                                  /    I'
                                 /    / \
                                /   VP   \
                              ???   |     \
                                    V'     \
                                   /        \
                                  CP         \
                                   |          \
                                   C'          \
                                  /  \          \
                                IP    \          \
                               / \     \          \
                              /   I'    \          \
                            VP    / \    \          \
                             |   /   \    \          \
                             V' /     \    \          \
                            /  /       \    \          \
      NP       NP      NP  V  I         C    V          I
      △        △        △
   Mary-ga  John-ga   hon-o  yon  da    to   omottei   ru
   Mary-SUBJ John-SUBJ book-DO read PAST that think    PRESENT
```

'Mary thinks that John read the book'