Superiority, Nesting and Crossing

1. ECP is not enough

We have learned an account of the contrast in (1)

(1) a. ??[Which book]$_1$ did you ask who bought t_1?
b. *[Which person]$_1$ did you what t_1 bought t?

We have provided an account of this contrast in a system that has two constraints against non-local movement (Subjacency which applies to all movement operations, and ECP, which only restricts the movement of subjects and adjuncts). However, there seems to be something that blocks (1)b independently of the ECP (as pointed out by Omer last week).

(2) a. ??[Which book]$_1$ did you ask who$_2$ Mary told t_2 [PRO to present t_1]?
b. *[Which person]$_1$ did you ask what$_2$ Mary told t_1 [PRO to present t_2]?

(3) a. This is the violin ~~wh~~$_1$ that I wonder which sonatas$_2$ to play t_2 on t_1.
b. *These are the sonatas ~~wh~~$_1$ that I wonder which violin to play t_2 on __.

2. Constraint on Crossing Dependencies (Kuno and Robinson)

The Constraint on Crossing Dependencies (CCD):
a. Two *wh*-dependencies cannot cross.
b. Two dependencies (chains) C and C' are called crossing dependencies if the head of C c-commands the head of C' and the tail of C c-commands the tail of C':
$C_{<head>}\ldots C'_{<head>}\ldots C_{<tail>}\ldots C'_{<tail>}$
c. Two dependencies (chains) C and C' are called nested dependencies if the head of C c-commands the head of C' and the tail of C' c-commands the tail of C:
$C_{<head>}\ldots C'_{<head>}\ldots C'_{<tail>}\ldots C_{<tail>}$

Frazier and Fodor (1978): The CCD follows from the nature of the parsing mechanisms that enable "fillers" to be associated with "gaps". Fillers are stored in memory by a "last-in-first-out" device (a "stack").

3. Superiority in English

Problem #1 (Superiority): We seem to be loosing a generalization

We might want to relate the contrast in (2) and (3) to the contrast in (4)

(4) a. You asked who$_1$ Mary told t_1 [PRO to present what].
b. *You asked what$_1$ Mary told who [PRO to present t_1].

Pesetsky (1982): (4)b involves an LF violation of the CCD (which Pesetsky generalized and called the path containment condition PCC)

(4') LFs of the sentences in (4):
a. You asked what$_2$ who$_1$ Mary told t_1 [PRO to present t_2].
b. * You asked who$_1$ what$_2$ Mary told t_1 [PRO to present t_2].

These LFs are predicted by the Extension Condition, which is needed on independent grounds, hence provide a very interesting unified account for (2), (3) and (4). Conversely, the facts in (2), (3), and (4) provide independent evidence for covert *wh*-movement. [To use the terminology of our class on covert movement, the CCD serves as a structure detector which indicates that there is covert movement.]

Question: What would one need to say in order to apply the Frazier and Fodor idea to account for an LF constraint against crossing dependencies?

4. Superiority in Bulgarian

Problem #2: Our generalization is wrong

There is evidence from Bulgarian against the CCD:

(5) a. Koj$_1$ kakvo$_2$ t_1 vižda t_2?
who what sees
cf. *Who sees what?*

Moreover, in Bulgarian crossing dependencies are preferred to nested dependencies:

(6) **Superiority Effect in Bulgarian (Rudin 1988)**
The leftmost *wh*-phrase in a Bulgarian multiple question is the *wh*-phrase that moves overtly in the corresponding English multiple question.

(7) a. Koj kakvo vižda?
who what sees
cf. *Who sees what?*

b.*Kakvo koj vižda?
what who sees
cf. **What does who see?*

(8) a. Koj k´de udari Ivan
who where hit Ivan
cf. *Who hit Ivan where?*

b.*K´de koj udari Ivan
cf. **Where did who hit Ivan?*

5. Richards's Proposal

Three components:

1. A derivational Theory of the effects we've seen in English: Attract Closest (Kitahara 1994, 1997, building on Kuno and Robinson 1972, Chomsky 1973, 1993, 1995)
2. Elimination of the strict cycle condition in favor of "featural cyclicity" (Chomsky 1995)
3. Tucking in (shortest move)[1]

5.1. Kuno and Robinson on Superiority in English

(9) An early statement of superiority
A *wh* word cannot be preposed crossing over another *wh*.
[Kuno and Robinson 1972]

This explains (4), but not (2-3):

(4) a. You asked who$_1$ Mary told t_1 [PRO to present what].
b. *You asked what$_1$ Mary told who [PRO to present t_1].

(2) a. ??[Which book]$_1$ did you ask who$_2$ Mary told t_2 [PRO to present t_1]?
b. *[Which person]$_1$ did you ask what$_2$ Mary told t_1 [PRO to present t_2]?

(3) a. This is the violin ~~wh~~$_1$ that I wonder which sonatas$_2$ to play t_2 on t_1.
b. *These are the sonatas ~~wh~~$_1$ that I wonder which violin to play t_2 on __.

5.2. Kitahara

Chomsky's account of superiority (4)

Attract closest: Every instance of *wh*-movement to C must be movement of the highest *wh*-phrase in the c-command domain of C.

Kitahara: this can also account for the PCC (2-3), if modified as follows: Every instance of *wh*-movement to C must involve movement of the closest moveable *wh*-phrase.[2]

5.3. Strict Cycle, the Extension Condition or Feature Cyclicity

Island conditions require a principle of cyclicity.

(10) **Extension Condition**: every instance of merge (internal, or external) must extend the structure.

[1] With a proposed unification with shortest move, which we will skip.
[2] This is slightly different from Kitahara's actual proposal, but will do for our purposes.

This condition would yield Bulgarian structures with the opposite order than that attested.

(11) **Feature Cyclicity**: If a head H needs to attract an XP, attraction must take place before any other operation.

Possible motivation: Late Merger.

5.4. Shortest Move

Consider a stage of the derivation of a multiple *wh*-question just before *wh*-movement takes place:

(12) C_{+wh}…Wh-phrase$_1$ vižda Wh-phrase$_2$?

At this point two things can happen: either wh-phrase$_1$ or wh-phrase$_2$ does. Attract closest determines that *wh*-phrase$_1$ moves before *wh*-phrase$_2$. This is the Chomsky-Kitahara explanation for English Superiority effects:

(13) Wh-phrase$_1$ C_{+wh} …t_1 vižda Wh-phrase$_2$?

If the Extension condition were postulated, we would get the wrong prediction for Bulgarian. However, if tucking-in derivations are allowed, the Bulgarian structure in (14) would be possible. Shortest move, insures that it is the only possible structure.

(14) Wh-phrase$_1$ Wh-phrase$_2$ C_{+wh} …t_1 vižda t_2?

5.5. New Prediction: A preference for crossing dependencies in Bulgarian.

Consider in greater detail the way Kitahara derives nested dependencies in English

(15) C_{+wh}…Wh-phrase$_1$ vižda Wh-phrase$_2$?

Shortest move determines that *wh*-phrase$_1$ moves to [Spec,CP]. Now another CP is constructed:

(16) C_{+wh} …Wh-phrase$_1$ C_{+wh}…t_1 vižda Wh-phrase$_2$?

At this point there is only one *wh*-phrase that can be moved. Movement results in a minor violation of subjacency (*wh*-island). The only way to derive a crossing dependency would involve a violation an early violation of attract closest.

However, that if *Wh*-phrase$_2$ were able to move to become a specifier of CP, we would predict the following (given shortest move):

(16') C_{+wh} …Wh-phrase$_1$ Wh-phrase$_2$ C_{+wh}…t_1 vižda t_2?

Which given attract closest would be transformed as follows to a crossing dependency

(17) Wh-phrase$_1$ C_{+wh} … Wh-phrase$_2$ t_1 C_{+wh}…t_1 vižda t_2?

Richards (2001) discovered that this is the attested pattern.

5.6. Evidence that the higher *wh*-phrase moves first (PMC)

Principle of Minimal Compliance: Only the first element that is the specifier of a X is subject to subjacency, shortest move, and attract closest.

Spell-out the predictions

5.7. Other constructions that show Bulgarain-type Superiority