Possible Ramifications for Superiority

1. Superiority up to semantic equivalence (Golan 1993)

 (1) Who knows what who bought?   (Lasnik and Saito 1992)
     Good but only when em **Attract Closest** bedded *who* receives matrix scope

Golan's Interpretation: *what* can move over *who* in the embedded clause because that is the only way for *who* to receive matrix scope.

Semantic Sensitive version of Attract Closest (SSSM): C_{uwh} must attract the closest *wh*-phrase it can *to derive a designated semantic interpretation*.

Can SSSM account for (the D-linking) exception to Superiority?

The answer would be yes if it turned out that superiority violations end up having different semantic interpretations (and if that difference depended on D-linking).

2. A possible source for different semantic interpretations

2.1. Dayal on uniqueness and its obviation in multiple *wh* questions

(2) Which boy came to the party?
    Presupposes: exactly one boy came to the party.
    #I know which boy came to the party; John and Bill came to the party.

(3) Which boy read which book?
    Two readings:
    Single Pair:
    I know which boy read which book; John read war and Peace.
    Pair List
    I know which boy read which book; John read war and Peace; Bill read BK…

2.2. An argument for a genuine ambiguity (restriction on Pair list):

When lists are given they must be exhausted: in other words, there are two options a single pair or an exhaustive list. No intermediate options are allowed.

 (4) a.   Guess which one of these 3 kids will sit on which of these 4 chairs
     b.   Guess which one of these 4 kids will sit on which of these 3 chairs
          (4)b is good only if there is a single pair as an answer (or if two kids will end up sitting on the same chair).

Conclusion: There is a genuine ambiguity. The pair list reading presupposes Exhaustivity (that every member of the set quantified over by the overtly moved *wh*-phrase is paired with a member of the set quantified over by the in-situ *wh*-phrase), a presupposition

which does not hold when a single pair serves as the answer.

Reading 1: presupposes that there is only a single pair as an answer to the question.
Reading 2: presupposes that there are exactly as many pairs and there are elements in the domain of quantification for the first *wh*P. (a consequence of exahustivity and point-wise uniqueness).

2.3. A possible source for different semantic interpretations

The nature of the exahaustivity and uniqueness presuppositions depends on which *wh*-phrase moves overtly. Hence, violation of superiority is predicted to yield different presupposition.

(5) a.   Guess which one of these 4 kids will sit on which one of these 3 chairs.
         Good only if there is a single pair as an answer.
    b.   Guess which one of these 3 chairs which one of these 4 kids will sit on.
         Good on a pair list answer.

(6) a.   *Except For War and Peace I know which boy read which book.
    b.   Except for War and Peace I know which book which boy read.

(7) a.   Except For John I know which boy read which book.
    b.   *Except for John I know which book which boy read.

(8) a.   I would like to know which resolution(s) Scott Brown voted for.
             (?)In fact, in every case I would like to know which senator voted for which resolution(s).
    b.   I would like to know which senator(s) voted for resolution 380.
             (*)In fact, in every case I would like to know which senator(s) voted for which resolution.
    c.   I would like to know which senator(s) voted for resolution 380.
             (?)In fact, in every case I would like to know which resolution which senator(s) voted for.

Prediction: violation of superiority should be impossible on a single pair reading.[1]

I think this might very well be right.

---

[1] This prediction seem to hold for German (Wiltschko 1997). Irene Heim pointed out an odd property of the current proposal, namely that it makes English and German out to be identical, contrary to what is usually assumed. If the proposal made here is correct, the only difference between English and German would be in the ease of getting a pair-list reading for multiple *wh*-questions (a reading that obviates superiority).

(9) My two kids are fighting:
    Who hit who first?
    *Who did who hit first?

(10)   a. I know which boy read which book. John read War and Peace
        b. I know which boy read which book. John read War and Peace, Bill read BK…

    c. *I know which book which boy read. War and Peace was read by John/John read
        W&P.
    d. I know which book which boy read. War and Peace was read by John, BK was
        read by Bill…

(11)   a. I know which one of my students wrote which one of these papers before any
            other student wrote any other paper.
        b. *I know which one of these papers which one of my students wrote before any
            other student wrote any other paper.

(12)   a. Guess which one of these 4 kids will sit on which one of these 3 chairs.
            Good only if there is a single pair as an answer.
        b. (#)Guess which one of these 4 chairs which one of these 3 kids will sit on.

(13)   a. A certain boy read a certain book. I wish I knew which boy read which book.
        b. A certain boy read a certain book. #I wish I knew which book which boy read.

(14)   a. There is a boy in that room reading a book. I wish I knew which boy is reading
            which book.
        b. There is a book in that room that the teacher is reading to a boy. I wish I knew
            which book the teacher is reading to which boy.
        c. There is a book in that room that a boy is reading. (#)I wish I knew which book
            which boy is reading.

3. A Golan type account of Superiority violations

Assume that wh_1 c-commands wh_2 prior to *wh*-movement

Moving wh_2 overtly over wh_1 is possible on the pair list reading because it yields
different presuppositions than what we would get if we moved wh1 overtly (hence
licensed by SSSM).

For concreteness, let's assume question have Hamblin denotations (i.e. they denote sets
of propositions or families of such sets, etc.).

We would now say that two question LFs Q and Q' are equivalent if
$\forall w \forall w'$ Ans($[\![Q]\!]$)(w')(w) = Ans($[\![Q']\!]$)(w')(w)

Superiority would state that moving wh_2 overtly over wh_1 is ruled out whenever the result is equivalent to what we would get by moving wh_1 overtly.

Possible approach to D-linking: Contextual assumptions must allow for two presuppositions to be met (*exhaustivity* and *point-wise uniqueness*). The universal exhaustivity presupposition can (realistically) be satisfied only if the domain that the overtly moved *wh* phrase quantifies over is somehow familiar (familiar enough to have presuppositions about). We thus derive something close to D-linking for the overtly moved *wh*-P.

Should the domain of the in-situ *wh*-P be familiar? Yes. Well at least enough for point-wise uniqueness to be grounded. [For related discussion, see Comorovski 1989, Kiss 1993, Miyagawa 2001, 2005, and Yoshida 2012.]

4. Problem: the constraint on crossing dependencies?

(15)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]?

(16)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 __.

It seems that SSSM should not block moving wh2 over wh1 in the embedded clause because that would yield a different interpretation globally.

5. Possible modifications

Strategy: keep to the standard (semantically insensitive) formulation of attract closest and introduce a semantically sensitive condition which will affect the consequences of attract closest in cases that involve two *wh* phrases in spec of the same CP.

5.1. Version One – with tucking-in

Internal merge (can sometimes not extend the tree…feature cyclicity)

Four Additional Assumptions

1.   Attract Closest receives the traditional formal definition, which is *semantically insensitive*.

(17)**Attract Closest:** C_{int} must attract the/a closest *wh*-phrase that has not already checked off a wh-feature on C.

2.   Shortest move -- the tucking-in requirement -- is semantically sensitive:

(18)**Semantically Sensitive Shortest move:** A moved *wh*-phrase must tuck-in below

another *wh* phrase, unless non-tucking in yields a different interpretation (at the end of the day).

3. Whether movement is overt or covert in English depends on the following principle:

(19)**Chain Pronounciation in English:** A head of a chain is pronounced if it is an outer-specifier of an extended projection. Otherwise the tail of the chain is pronounced. (Pesetsky 2001, see also F&P 2010 handout
http://web.mit.edu/linguistics/people/faculty/fox/Ben-Gurion-7-09.pdf).

4. If a chain is pronounced in a trace position, further movement cannot change the verdict (i.e. overt wh-movement of covertly moved wh phrases is impossible).

5.2. Going Through the various cases:

Let's start with a C_{int} that has two *wh*Ps in its scope and must attract at least one (i.e. is interpreted as a question and has an uninterpretable wh feature).

C will attract the higher whP (wh_1) by attract closest[2]

$wh_1 C_{uwh}…wh_1… wh_2$

Then three possibilities need to be considered

1. Do nothing:
$wh_1 C …\text{wh}_1… wh_2$ (lower position deleted)

2. Tuck-in wh_2 (allowed by SSSM, no matter what)
$wh_1 \text{wh}_2 C …\text{wh}_1… wh_2$ (lower position is deleted at PF for wh_1; higher position is deleted for wh_2)

3. Extend tree with wh_2 (allowed by SSSM if it yields a different interpretation than what could be derived from 2)
$wh_2 \text{wh}_1 C …wh_1…\text{wh}_2$ (lower position is deleted for wh_2 higher position is deleted for wh_1)

3 is what we call a superiority violation and it is allowed if it receives an interpretation that cannot be derived by 2. Two Scenarios to consider

#1 The CP we are constructing is interpreted as a multiple *wh* question. 3 and 2 would have different interpretations only if we have a pair list interpretation for both (which, in turn, requires the two whPs to be D-linked). The two pair list interpretations as we have seen are different.

#2 The CP we are constructing is interpreted as a single *wh* question. It is embedded

---

[2] If a highest *wh*P exists. Otherwise more than one *wh*P is a legitimate attractee.

under another CP and one of the wh phrases has wider scope. A different interpretation will be made possible by 3 only if wh_2 stays where it is and wh_1 has wider scope. But wh_1 cannot move overtly from Spec CP because there are no phonological features there – this position has been deleted (assumption 4). Hence it must move covertly as in L&S's example in (1).

5.2. Version 2 – no tucking in (based on conversations with David)

Merge (internal and external) must extend the tree

Four Additional Assumptions

1. Attract Closest receives a new definition, which cares about only one instance of movement, and in addition is *semantically insensitive*.

(20) **Attract Closest:** C_{int} must attract ***one*** *wh-phrase* which is closer to it than any other phrase.

2. Apology principle -- the tucking-in requirement -- is semantically sensitive:

(21) **Semantically Sensitive Apology:** When C_{int} attracts multiple *wh*-phrase its outer specifier (the last instance of movement) must be the closest attractee, unless a different choice yields a different interpretation.

3. Whether movement is overt or covert in English depends on the following principle:

(22) **Chain Pronounciation in English:** A head of a chain is pronounced if it is an outer-specifier of an extended projection. Otherwise the tail of the chain is pronounced.

4. If a chain is pronounced in a trace position, further movement cannot change the verdict (i.e. overt wh-movement of covertly moved wh phrases is impossible).

6. Superiority and Pair list readings

(23)   a. What did every boy read?   (pair list available)
        b. Who read every book?   (pair list unavailable)

Could this follow from superiority?
        Yes, if the condition has no exception for different semantic interpretation: A-bar movement to the extended projection of C must preserve c-command relations from A positions.

Possible Prediction (based on the Bulgarian pattern):

(24)   a. I know which girl you told a story about Plato, and which girl you told a story about Socrates

       b. #In fact I know which girl you told a story about EVERY philosopher.

       a. I know which girl Bill told a story about Plato, and which girl Fred told a story about Socrates
       b. (#) In fact I know which girl EVERY boy told a story about EVERY philosopher.

Assuming we want Superiority to derive the constraints on pair list readings, what do we do?

7. Third Version (Pesetsky 2001 Meets Golan…based on Conversation with Hadas Kotek)

Superiority is not semantically sensitive.

Possible implementation (always extending tree)

1. Attract Closest receives a new definition, which cares about only one instance of movement, and in addition is *semantically insensitive*.

(25)**Attract Closest:** C_{int} must attract *one wh-phrase* which is closer to it than any other wh-phrase

2. Apology principle -- the tucking-in requirement -- is also semantically insensitive:

(26)**Semantically insensitive Apology:** When C_{int} attracts multiple *wh*-phrase its outer specifier (the last instance of movement) must be the closest attractee~~, unless a different choice yields a different interpretation.~~

What is semantically sensitive is the principle that tells us which apparent wh-phrases are real wh-phrases

3. Every *wh* phrase can be marked as a non-candidate for movement if this marking yields a different semantic interpretation.

4. Whether movement is overt or covert in English depends on the following principle:

(27)**Chain Pronounciation in English:** A head of a chain is pronounced if it is an outer-specifier of an extended projection. Otherwise the tail of the chain is pronounced. (cf. F&P 2010 handout http://web.mit.edu/linguistics/people/faculty/fox/Ben-Gurion-7-09.pdf).

5. If a chain is pronounced in a trace position, further movement cannot change the verdict (i.e. overt wh-movement of covertly moved wh phrases is impossible).