Bulgarian Superiority and Minimalist Movement Theory

John Frederick Bailyn
Stony Brook University

1   The Bulgarian Superiority Puzzle

As everyone who has taught introductory syntax knows, English Superiority of the kind found in (1)a)-(1)b) vindicates Minimalist claims about architecture and economy.

• Superiority (English-style):

a)   a. **Who** said **what**? $SUBJ_{wh} > OBJ_{wh}$
    b.***What** did **who** say?   *$OBJ_{wh} > SUBJ_{wh}$

b)   a. **Who** works **where**? $SUBJ_{wh} > LOC_{wh}$
    b.***Where** does **who** work?   *$LOC_{wh} > SUBJ_{wh}$

Minimalism's bottom-up system of syntactic construction (generally know as Bare Phrase Structure, see Chomsky 1995, a.o.) is typically accompanied by top-down triggering of movement, in that a Probe X (with relevant feature F) searches within its c-command domain for a matching feature on a Goal Y with which it establishes an Agree relation. Overt movement is then parasitic on Agree, given the proper specification of a sub-feature of [F] (its "strength", or EPP feature) which forces Y to move so that the eventual relation between X and Y be maximally local. Together with an Economy principle such as Attract Closest (see (2)), this system allows effects such as English Superiority to be handled in a highly elegant fashion, in that the structurally highest WH (here the subject) is always the one to be attracted, because it is

closer to the Probe, hierarchically speaking, than some Z that is lower in the structure such as the object WH. (a)-(b)b are underivable.

2) Attract Closest (descriptive): If an attracting head X (a Probe) carries a (strong) feature [F] and two elements Y and Z (Goals) also carry [F] then the **closest** [F] element wins (closest defined by c-command) (Chomsky 1995, Richards 1997, Pesetsky 2000)

This account does away with any need for the ECP, Proper Government, Comp indexing or other devices required in theories such as Government and Binding (Chomsky 1995, ch. 1). Just pure Probe-Goal and Economy.

However, under Minimalism things don't work out so nicely for the Superiority effects shown in (3) from a multiple overt WH-movement language such as Bulgarian:

• Superiority (Bulgarian-style): (Rudin 1988, Bošković 1998, Grewendorf 2001)

3) a. **Koj**   **kogo**   e   vidjal? $SUBJ_{wh} > OBJ_{wh}$
        who   whom   aux   seen   (Bulgarian)
        "Who saw whom?"

    b. *__Kogo__   **koj**   e   vidjal? $*OBJ_{wh} > SUBJ_{wh}$
        whom   who$_{NOM}$ aux   seen   **Superiority**
        *"Whom did who see?"

All WH-elements must front in Bulgarian. It has been known since at least Rudin 1988 that after multiple WH movement, the originally structurally higher element must precede the structurally lower one. Further, the WHs form a single constituent that precedes auxiliaries and parentheticals, as shown in (4).

4) a. **Koj** **kakvo** ti e   kazal?   [WH-1, WH-2, WH-3] > clitics
       who   what   you aux   told
       "Who told you what?"

    b. *__Koj__ ti e   **kakvo**   kazal?   *WH-1 > clitics > WH-2, WH-3
       who you aux what   told
       "Who told you what?"

At the same time, it has been claimed that non-structural factors play a role in determining better and worse orderings among Bulgarian WHs, especially   animacy,   topicality/discourse-salience,   and   type   of

adjunct/modifier. This had led some to deny the structural basis of Bulgarian Superiority (as in Jaeger 2004).

However, the subtle contrasts among orderings in Bulgarian Multiple WH constructions are examined in detail in Krapova & Cinque 2005, whose cartographic approach could be expected to challenge claims of uniform landing sites competing for similar Goals and who propose distinct landing sites for various kinds of WH elements. It is all the more persuasive, therefore, that Krapova & Cinque conclude that although there are nuances among various orders, especially with adverbial WHs, subjects always precede other elements, and surface order generally mimics base order: "the surface order of Bulgarian … wh-phrases … appears to reflect *their relative order prior to wh-movement*" (Krapova & Cinque 2005: 190, emphasis mine) Krapova and Cinque conclude that a generalized principle of structure preservation is required, that is, some kind of principle that guarantees Superiority is respected, even in this most articulated cartographic approach.[1] [2] So we can be confident that there is some sort of Superiority effect in Bulgarian to be accounted for, one which has the effect of structure preservation. I will now turn to how minimalist accounts of Bulgarian Superiority generally run.

2   The Standard Account

The standard account rests on several assumptions. One preliminary assumption, needed by everyone, it appears, is that some kind of *lexical property* of Bulgarian WH-phrases requires them to move:

• Assumption 1 (needed for all): Bulgarian WHs have a property that *requires* them to move

Here's how the standard story goes: we have WH-1, a subject, and WH-2, an object, let us say. Standard top-down Probe-motivated movement is assumed, as it is for English.

---

[1] I will not address their own version of this principle, deriving from feature-based Relativized Minimality (of the kind in Rizzi 2004), because they term it "highly tentative" and it does not appear to be fully worked out or empirically adequate.

[2] I do not address the question of triple WH-questions since there is no consensus on the ordering restrictions (or lack thereof) within them.

STEP 1: Structure is built up to the level of C. C probes down and finds the closest relevant element, WH-1, as it does in English, pulling it up to a local Specifier position. Here two additional assumptions come in: First, it is assumed that Bulgarian C allows multiple specifiers. Second, C is also claimed to be distinct in Bulgarian in that it continues to Probe after the initial WH-1 is located, matched with and pulled up. This is the famous Probe-that-Keeps-on-Probing.

• Assumptions 2-3: Bulgarian has a "special" kind of C_{wh}: ("The Probe-that-Keeps-on-Probing"

    (i) it allows multiple specifiers and
    (ii) it allows non-deletion of its uninterpretable feature (to attract more WH_s)

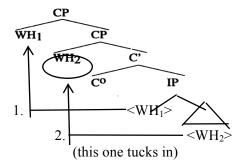**STEP 2:** C thus probes on, finds WH-2, and agrees with it.

• Assumption 4: There can be "Tucking-In" (movement to a lower Specifier)

STEP 3: C pulls WH-2 up as well, and, crucially, "tucks it in" to a lower specifier position (Richards 1997, 1999 a.o.) The *necessity* that WH-2 land in the Tucked-in landing site is motivated by an additional Economy principle – Shortest Move, which requires that WH-2 move to the *closest* available position. An additional assumption is thus required here, namely that the lower Spec is crucially closer to the base position than the outer Spec, so that Shortest Move forces WH-2 to tuck in.

• Assumption 5: Tucked-in elements, in lower Specs, are not equidistant. Shortest Move requires movement to this "closer" Specifier for WH_2

Underlying order (WH-1 > WH-2) is thus preserved. The order of operations is shown in (5).

5) Schematic picture of Bulgarian Superiority obeying derivations such as (3)a)

In what follows, I first present what I consider to be serious problems with this standard account of Bulgarian superiority. Then, I propose an alternative that solves those problems, which, though admittedly unusual, in fact solves the problems I identify in a principled fashion, without machinery that is not already independently needed in the grammar. Finally, I present two welcome consequences of the account.

3   Problems with the Standard Account

So what's wrong with the standard account? In fact, there are quite a few problems with it, some conceptual some empirical. In what follows, I will discuss five major issues with the standard account.

Problem 1. The primary concern, at least for those working within the architectural framework of Bare Phrase Structure (Chomsky 1995), is that *Tucking-In is counter-cyclic, violating the Extension Condition*, as loosely given in (6): (Adger 2003: 95)

6) **The Extension Condition**: A syntactic derivation can only be continued by applying operations to the root projection of the tree.

In my view it is serious business to abandon Extension, the core property of Bare Phrase Structure, and the property that allowed us originally to dispense with X'-theoretical templates (or Phase-Structure rules) and the central component needed for accurate labeling as well, if Chomsky's 2013 approach to the matter is on the right track. To ask us to abandon Extension to allow Tucking-in is far less innocuous than those advocating the standard approach to Bulgarian Superiority would have us

believe. It is it not at all clear how Bare Phrase Structure can function without some such principle of Extension; at very least proponents of Tucking-In have not, as far as I know, addressed the issue of projection and labeling without Extension.[3]

---

[3] Without answering the issue of how BPS can function without Extension, Richards (1999) does provide an empirical argument for Tucking-In, relying on The Principle of Minimal Compliance:

i) The Principle of Minimal Compliance (PMC) (Richards 1997, 1998, 1999) (loosely):

> "permits the computational system to "ignore" those portions of a syntactic structure that have already been determined to be participating in a well-formed dependency" (Richards 1999: 137)

Bulgarian Subjacency violations such as (iia) are thus said to be improved due to a higher instance of Subjacency-obeying WH-movement, as in (iib):

(ii)   a. *[**Koja kniga**]$_2$ otreče   senatorât   [mâlvata   če
            which book   denied   senator.the   rumor.the that

        [pravitelstvoto   iska   da   zabrani **t$_2$**]]?
        government.the   wanted to   ban

"Which book did the senator deny the rumor that the government wanted to ban?"

    b. ?[**Koj   senator**]$_1$ [**koja   kniga**]$_2$ t$_1$ otreče   [mâlvata   če
        which   senator   which   book   denied   rumor.the   that

        [pravitelstvoto   iska   da   zabrani **t$_2$**]]?
        government.the   wanted   to   ban

"Which senator denied the rumor that the government wanted to ban which book?"

Assuming the PMC, and that the legitimate move must precede the potentially illegitimate one, Richards (1998, 1999) derives an argument that WH$_2$ must move second, and therefore tuck-in, to derive the proper surface order. Richards also shows that multiple WHs starting from non-c-commanding positions have to obey a certain linear order in Bulgarian, consistent with the PMC. However, the nature of the amelioration effect is unclear – Grewendorf (2001) shows it does not apply to certain violations (adjunct islands etc). Thus the PMC is generally suspect. And without the PMC, the facts in (ii), along with others in Richards 1997, 1999, do not constitute an argument for the *orderings* of multiple movements. The apparent correct generalization is that the acceptable surface order of WHs that result from multiple WH-movement always shows structure preservation. These facts are consistent with the approach here, whereby WHs move from the bottom up, forming structure-preserving clusters as they go.

Problem 2. Not only must the standard account of Bulgarian Superiority allow multiple specifiers, an assumption that those working in cartographic frameworks do not allow on independent grounds, it crucially *must assume that multiple specifiers are not equidistant from the rest of the tree*. (Otherwise, Shortest Move would not *force* Tucking-In, and the object would at least have the option of not tucking–in, deriving an ungrammatical order.) The idea that multiple specifiers are not equidistant contradicts successful analyses of various multiple Spec constructions such as Grewendorf and Sabel 1999's contrastive analysis of German vs Japanese Long Distance scrambling, and various others. Non-equidistance of multiple Specs in some cases but not in others could of course be justified with the right kind of fancy maneuvering, but what I would like to suggest is that we'd all be better off in a world where if there are multiple Specs, they are at very least equidistant from the rest of the tree. If one's basic assumptions allow multiple Specs, as mine do. they need to be equidistant. If multiple Specs are not permitted in principle (as with Cartography), then you need to use Specifiers of distinct categories, with the lower WH in the Spec of a lower category, rendering the Extension violation even more problematic, and the constituency of the Bulgarian WHs more difficult to account for. So everyone should be somewhere between somewhat unhappy and very unhappy with the current claim of non-equidistant multiple Specifiers.

Problem 3. Two different economy conditions are needed for Bulgarian superiority - Attract Closest and Shortest Move.

Problem 4. We need to posit the Probe-that-Keeps-on-Probing. Though possibly collapsible with multiple Specifier projection, the Keeps-On-Probing quality of the Bulgarian C head involves a kind of Extremely Enlightened Self-interest (the Probe must continue to probe for exactly as long as is needed to pull up all the WH's present), and be satisfied with exactly no fewer and no more. C is thus omniscient, and its requirement have at least a highly descriptive (if not an entirely fortuitous) character. This is of course undesirable.

Problem 5. Finally, there are *three* claimed parameterized distinctions between English and Bulgarian. First, there is the one difference any theory needs to acknowledge, namely that (i) Bulgarian WH elements differ from their English counterparts in having to move. However, the

standard account requires positing two further differences: (ii) that the Bulgarian C head can have multiple Specifiers, and (iii), that the Bulgarian C head is a Probe-that-Keeps-on-Probing.

(7) summarizes the problems identified with the standard analysis:

(7) **<u>Problems with the standard analysis</u>**:
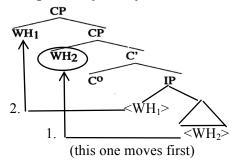
  A. Tucking-In is counter-cyclic (violates the Extension Condition (6), and is is unmotivated by Economy)

  B. Multiple specifiers must crucially *not* be equidistant, contrary to both evidence (Richards 1999, Grewendorf & Sabel 1999, Grewendorf 2001) and intuition, in order for Shortest Move to force Tucking-In.

  C. *Two* economy conditions are required (Attract Closest and Shortest Move) for Bulgarian Superiority

  D. The account needs the Probe-that-Keeps-on-Probing

  E. Parametrization is required in three (related) aspects:
      (i) in the lexicon (Bulgarian WH elements differ from their English equivalents in having to move)
      (ii) in the ability/need to have multiple Specifiers (Bulgarian) vs a single Specifier (Eng).
      (iii) in the nature of the [wh] features of C (it must be a multiple rather than single attractor)

Before turning to my alterative, I should acknowledge other accounts that deserve mention: Rudin's original 1988 account, whereby WH-2 adjoins directly to WH-1, but to its right (the latter being a required but undesirable stipulation), and Grewendorf's 2001 account that allows WH's to cluster in the course of the derivation, though also with a stipulation of right-edge landing site. (The right adjunction/landing is their equivalent of Tucking-In.) My primary goal is to find a way to avoid this as well, without introducing extra machinery or assumptions. The rest of this article will be devoted to that attempt.

4 A Possible Solution

The way to do that, of course, is to abandon the assumption of top-down triggering of movement in such derivations. As Richards himself admits: "If we wish to maintain the idea that movement always expands the tree, creating a Specifier higher than all the existing structure, we must apparently conclude that the lower of the two WH-words must move first" (Richards 1999). Top-down Probe systems don't allow that, so Richards pursues Tucking-In. But imagine, instead, that WH-2 somehow *is* able to move first, ending up extending the tree, intuitively speaking, after which WH-1 moves above the initially derived structure, and original order is preserved: WH-1 > WH-2. That's the angle I am pursuing here. This is shown schematically in (8).

8) Picture of Bulgarian Superiority if WH-2 moves first (no Tucking-In)

(this one moves first)

Here's the main idea: WH-2 *does* move first, so long as Goal-driven (or "self-motivated") movement occurs in such cases. How might this come about? Here, I follow the primary claim of Bošković 2007 ("On the Locality and Motivation of Move and Agree: An Even More Minimal Theory"). In that article, Bošković assumes movement is (*always*) triggered by a morphological requirement of the moved element itself (hence the term Goal-driven movement, which I call Self-motivation, a version of the old Greed principle). In my approach, such elements are equipped with what I call a Blinking Blue Light, which is the equivalent of Bošković's [uK] feature. "In a multiple *wh*-fronting language like Bulgarian, *wh*-phrases would be obligatorily specified with a uK feature" (Bošković 2007):

9) Self-motivated movement: (Bošković 2007: 609)

$$[_{XP} \ldots X \ldots \quad Y]$$
$$i\text{F}$$
$$u\text{K} \ [\leftarrow \text{the Blinking Blue Light}]$$

Bošković's [uK] feature is essentially an instantiation of the need to move as soon as possible. Their need to move *is* the Blinking Blue Light, this is the source of such WH elements (always) undergoing self-motivated movement. Conversely, *everything* that undergoes self-motivated movement (and *only* those things), are elements with a similar feature, lexically determined.[4]

It is important to note that although Bošković's focus is English WH-elements undergoing successive cyclic Long-distance movement, which I will in the end argue *do not* undergo self-motivated movement, the core idea here comes from there. For me, the Blinking Blue Light property of Bulgarian WH elements follows as an instantiation of the shared assumption everyone needs about Bulgarian WH's, namely that they have a distinct lexical property that forces them to move.

The natural question arises for any claim of self-motivated movement, HOW DO SELF-MOTIVATORS MOVE BEFORE THEIR EVENTUAL PROBE IS PRESENT IN THE STRUCTURE? Bošković 2007 does not fully answer this question, other than to assume that movement to phase edges is generally possible as an option, which, if not taken, leads to failure associated with the higher Probe when the time comes to close off a phase (basically a bottom-up version of the Phase Impenetrability Condition). The elements cry "I need to move", but where can they go? There is no feature-driven position available to go to.

Here, then, we get to the technical core of my proposal, – elements with Blinking Blue Light undergo *immediate* movement of the kind that has been called **Sidewards Movement** (see 10)), (Nunes 2001, 2004), exiting their subtree and immediately merging with C, the element they share the relevant [wh] feature with.

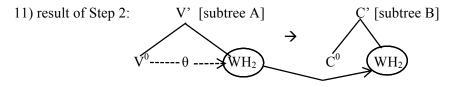10) Sidewards Movement (Nunes 2001, 2004)

---

[4] I depart from Bošković's 2007 overall approach in one crucial respect - for Bošković this feature is found on all moving elements; here, it is only found on obligatorily moving elements, a possibly more intuitive and successful assumption, to be explored below.

- (sub)trees are built in parallel in the "workspace"

- Sidewards movement allows an element in a partially built structure to dis-attach and move "sidewards" into a distinct subtree. The theory allows this freely, subject to linearization.

Other than Sidewards Movement (an option Bare Phrase Structure certainly allows, as Nunes (2001, 2004) argues convincingly), I assume nothing other than bottom-up derivations, the Extension Condition, and that Specifiers precede Complements (or, more exactly, that 2nd merge elements precede 1st merge elements). Here's how the derivation works:

STEP 1: WH_2 is merged in base position in the usual way (for theta-assignment) in Subtree A).[5]

STEP 2: WH_2 has uK (the Blinking Blue Light). This forces it to move sidewards and join with C^0 [wh] (already in the Numeration/Workspace), creating Subtree B and shown in (11):

11) result of Step 2:　　　V' [subtree A]　　　　　　C' [subtree B]
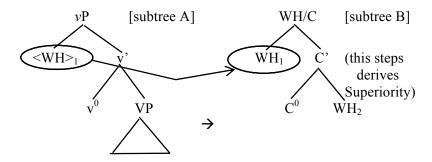
V^0 ------ θ ----→ WH$_2$　　　C^0　　WH$_2$

→

Assuming standard right-branching LCA-driven principles of combination, the first Merge position is to the right of the relevant head, and the second Merge is to a higher c-commanding Specifier type position to the left, (which is where WH-1 will eventually Merge, deriving the required surface order). The derivation proceeds.

STEP 3: WH_1 is merged into its base position in the usual way.

---

[5] For A-movements necessary to establish case and other relations, I assume that Sidewards Movement launches from theta position (as a function of the Blinking Blue Light), and that case checking (and any A-movement involved) involves the remaining silent copy of the WH phrase. With subject Wh phrases, the same would apply at Step 3, with regard to movement to SpecTP for EPP purposes. Thanks to an anonymous reviewer for suggesting this possibility.

STEP 4: WH_1 has uK (the Blinking Blue Light). This forces it to move sidewards, merging with the already existing C+WH-2 cluster (Subtree B), as shown in (12):

12) Result of Step 4:[6]

In terms of the WH elements, this is similar to Grewendorf's 2001 cluster account, quoted here (except that his account requires WHs attracting other WHs, and right-attachment of rising WHs). He says:

> I wish to suggest an analysis of overt multiple wh fronting that does not proceed from the idea that wh-elements move individually to [Spec, CP]. I will argue that it is instead *a cluster of wh-elements* that moves to [Spec, CP] in languages such as Bulgarian, and that assuming the formation of a wh-cluster prior to wh-movement to [Spec, CP] avoids the conceptual and empirical problems of the other approaches mentioned here and enables us to answer the question of what motivates multiple wh-fronting.
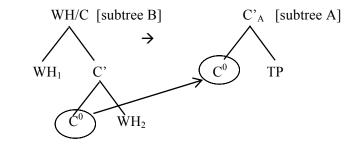
My approach obviously shares this aspect of Grewendorf's. However, it does not require any right attachment. The derivation proceeds.

STEP 5. At the moment when TP is complete, we see an unusual, but entirely plausible, case of what I call Backwards Sidewards Movement,
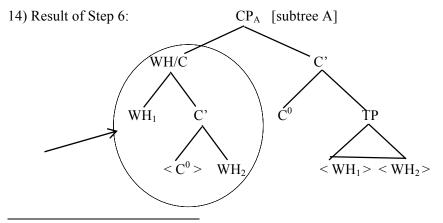
---

[6] A reviewer raises the issue of why the clusters resulting from this step do not form CED islands. After Step 5, when they are headless clusters, and without whatever restricts movement out of true conjunctions, I see no reason why they should resist further extraction (in fact, as we know from analyses of Polish-style multiple Wh constructions, further movements after similar clustering is standardly assumed, see below).

whereby C exits Subtree B, (leaving a "bare" WH cluster behind), and joins Subtree A in the manner C always does, namely merging with TP, as shown in (13). This movement is driven by the standard c-selectional needs of C and the requirement that it end up in the proper place (above T), in the "hierarchy of projections" (Adger 2003).

13) Result of Step 5:

STEP 6. We now integrate the two trees, by merging Subtree B (the WH cluster) with the top of Subtree A (the existing C projection (call it C' or CP as you like). This is shown in (14):[7]

14) Result of Step 6:

---

[7] A reviewer raises the issue of the motivation for Step 6, noting that in principle the move is not required for feature checking of the [+wh] elements in Subtree B. However, we can assume that something akin to the original wh-criterion (Lasnik & Saito 1992) namely that [+wh] elements must appear in a [+wh] CP, will require it. Alternatively, it could be an optional move, in terms of features, but a derivation without it will never be integrated, and the Sidewards movement never resolved, the copies never deleted, etc. I leave the matter open.

Constituency is now as we need it to be to account the constituency facts identified by Rudin 1988 and others for Bulgarian type languages, and we share Grewendorf's advantage over other accounts that no Multiple Specs are needed, no Probes-that-Keep-on-Probing are needed, and, crucially, no Tucking-In is needed. The derivation converges. A summary of the resulting theory of movement is provided in (15):

15) Summary of the hybrid proposed theory of movement:

  a. Single movements (Attract) are driven by a (strong) feature of the Probe (standard top-down assumption)

  b. Multiple movements to a single head are driven by a (strong) feature [uK] *of the moved element itself* (cf Bošković 2007) (this is Self-motivated Movement)

  c. Derivations are entirely bottom up. Self-motivated movement begins before the checking head (Probe) is merged (=Bošković's "early" movement)

  d. Sidewards movement always takes place with Self-motivated Movement

A note on linearization: For the WHs, I assume that the cluster shares elements of both of its WH conjuncts, as in a ConjP structure. The cluster in SpecCP, sharing both WH-1 and WH-2 features, then c-commands the lower WH copies of each, and the WH elements are linearizable. As for C, either we have to claim that a head c-commands into its Specifier (not an unreasonable assumption, given that the Spec is the position to which overtly raised element are located), and so the C in the cluster is deleted, as desired, or we are left with a kind of Remnant Movement structure, whose gap needs to be analyzed on a par with other cases of remnant movement. The fact that this structure is created by 2 instances of Sidewards movement does not matter in that regard.[8]

The account has obvious advantages, summarized in (16) below.

---

[8] It is also possible that Sidewards Movement is not copy-based (even if other movements are) and only reintegration is required, not linearization of Chains as in the original Nunes view of things. If so, then the issue of linearization of C does not arise.

16) Summary of the account's advantages:

A. It derives structure preservation in Bulgarian Superiority instances in a principled fashion. Bulgarian-Superiority is pure Structure Preservation - a side effect of the system's architecture; nothing special needs to be assumed (other than BPS – Merge, Extend, Copy, etc.)

B. It dispenses with multiply active Probes: all multiple moving elements are driven by their own strong features. The Probe probes once. (It also raises the questions of whether all purported multiple Spec instances can be derived in the same manner, though limited of course to those checking with a single head).[9]

C. It dispenses with the assumption of Probes-that-Keep-on-Probing.

D. It derives the right word order without resorting to Tucking-In.

E. The various WH elements are equidistant (for purposes of further movement).

F. Parameterization reduces to the one lexical difference that everyone needs: – the (lexical) property of Bulgarian WH phrases requiring them to move (here the Blinking Blue Light). This property itself triggers Sidewards Movement and cluster formation – everything else results from the independently needed workings of the system.

G. Finally, there is only one relevant Economy Principle for both kinds of Superiority – Attract Closest. Shortest Move is not needed.

5 Extensions

There are two positive extensions of significance: the first involves the resulting analysis of Russian/BCS/Polish type languages, in which all WHs move but which apparently lack Superiority (see (17)), and which show distinct properties from Bulgarian (see list in (18)). The second set of consequences concern locality. I turn to these now.

17) a. **Ko koga** vidi? b. **Koga ko** vidi? (BCS)
who$_{NOM}$ whom$_{ACC}$ sees   whom$_{ACC}$ who$_{NOM}$ sees
'Who sees whom?'   'Whom does who see?

---

[9] Note that the account does not dispense with multiple Specifier constructions in general – it just derives them through Sidewards movement, and maintains the assumption that they are equidistant for the purpose of further movement.

18) Properties of BCS/Russian type languages (Rudin's (1988) [-MFS] languages):
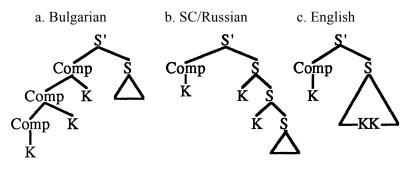
a. **+MFS languages**:   b**. –MFS languages**:
   (Bulgarian / Romanian)   (SC/Polish/Russian)

   i. Parentheticals/ clitics come   i. Parentheticals/ clitics come
      after all WHs   after 1st WH
   ii. multiple WH extraction possible   ii. multiple WH-extraction not possible
   iii. Superiority holds   iii. Superiority doesn't hold

5.1   Russian/SC apparent lack of superiority
For the Russian/Polish/SC type languages, I follow Rudin's (and everyone else's) general idea about constituency of WH's in these languages, (schematized in (19)b):

19) Rudin's (1988) Wh movement types: (K] here = [wh])

a. Bulgarian   b. SC/Russian   c. English

It is generally assumed, following Rudin 1988, that all but one of the WH elements cluster at the TP edge; the leftmost one is in SpecCP (see also Bošković(1998, 2002 etc). The position of clitics and parentheticals after WH-1 is thus accounted for (basic placement at the two distinct edges for Bulgarian vs SC/Polish is maintained in this account).

Standard accounts assume a TP-level feature driving non-wh-movement of WH elements in the BCS/Polish-type languages, usually Focus, though I take no stand on that particular feature. Regardless of the driving feature. the question remains: *why should Superiority not*

hold in such languages? Some accounts are given in Rudin 1988, Bošković 1997, Richards 1997, and Stepanov & Stateva 2009.

For Rudin (1988), the workings of the Empty Category Principle conspire to allow (17)b, an approach no longer available under modern assumptions. For Bošković 1997, single WH-mvt happens first, followed *later* by *lower* adjunction to IP, an undesirable counter-cyclic stipulation. ("I leave it open here how this should be reconciled with the cycle" Bošković 1997: 12). For Richards 1997, "IP absorption" is not subject to Superiority but "CP absorption" is, a distinction that is not further accounted for. Standard accounts assume that multiple movement driven by Focus are simply not subject to Superiority whereas multiple ("true") WH-movements are. Superiority is claimed not to apply because the inherent [Foc] movement requirement of Rus/BCS WHs forces the movement, so each element is driven separately (so no competition): "Attract/Shortest is simply irrelevant… Each *wh*-phrase in a multiple *wh*-question moves for an independent reason [Focus]" (Stepanov & Stateva 2009, following Stepanov 1998). In addition to its stipulative nature, such an approach obviously cannot apply here, given the architecture described above for self-motivated WH-movement since the Focus movement to the TP edge is also self-motivated, being obligatory for all WH element sin these languages.

A more useful account runs as follows: WHs in these languages are also Self-motivated movers, as in standard accounts -- their obligatory movement for Focus purposes is instantiated as their having what we could call a Blinking Purple Light, and therefore they also undergo Sidewards Movement just as we have seen for Bulgarian WHs. (Note that the Blinking Purple Light is no different from the standard assumption that these are "TP-absorption" languages as in Richards sense.) These WHs form a (Superiority-obeying) [Foc] cluster in a Subtree in parallel fashion to what we have just seen for Bulgarian [wh]. Superiority *appears* not to hold only because of *subsequent* movement of one element to a higher spot. This final move is standard Probe-driven *single* WH-movement (this piece is based on Scott 2012).[10]

---

[10] This has the additional advantage of rendering all overt WH-mvt languages the same in that the feature movement of their C elements are the same– a singly probing Probe with a strong (EPP) WH feature attracting a single WH element (which could be a cluster) to a single Specifier.

The choice of WH element to undergo this final move is free, due to the natural (and expected) equidistance of elements in the TP level cluster to a higher Probe.[11] Crucially, the cluster *does* obey Superiority, an effect identified by Scott (2012) as The Emergence of Superiority.[12] When the additional movement is blocked, Superiority emerges as predicted, as shown in (20)-(21) for BCS and throughout Scott's dissertation for Russian.

Emergence of Superiority in BCS: (Bošković 1997, 2002, Stjepanović 1999)

• subordinate clauses:

20) a. Jovan i Marko ne znaju **ko** je **koga** istukao.
   Jovan and Marko not know who is whom beaten
   'Jovan and Marko do not know who beat whom.'

   b. *Jovan i Marko ne znaju **koga** je **ko** istukao.
   Jovan and Marko not know who is whom beaten
   *'Jovan and Marko do not know whom who beat.'

• overt topics:

21) a. Tom čoveku, **ko** je **šta** poklonio?
   that man who is what bestowed

   b. ??Tom čoveku, **šta** je **ko** poklonio?
   that man what is who bestowed

No parameterization of Superiority is required, or permitted, or ever possible (marking the end, by the way, for both Bruening's (2001) account of Scope Freezing and Stepanov and Stateva's (2009) attempt to

---

[11] This assumption incidentally renders the Tucking-In account additionally untenable, since multiple Specs must crucially *not* be equidistant for that account to work, in that the lower of the Multiple Specs must be closer to the lower part of the tree than the higher one to force Tucking-In to give the correct surface order in Bulgarian. The Richards-style accounts can not, then, derive lack of Superiority in the Russian/SC type languages in this manner. And in fact, it is never explained why CP-absorption obeys superiority but IP-absorption does not. For me, they both do, but a single further move obscures it in the Russian/SC case.

[12] Note that I differ from Scott in associating Russian with the BCS/Polish-type languages and not with Bulgarian.

correlate WH and QR Superiority, both of which have been shown to encounter major problems anyway, see Antonyuk 2015). The principles that lead to Superiority of this kind are independent parts of the system (Merge order and bottom-up derivations) and of course would not be expected to be parameterized. This is obviously a welcome result.

5.2   Locality
The second consequence involves Locality. Here I assume a version of the Bošković 2007 story on successive cyclicity, namely a *non-feature checking* account of intermediate stopping points in standard successive-cyclic LD WH-movement. Sidewards movement, however, *requires* feature checking for reintegration into Subtree A.   Thus any WHs being attracted by main clause C as part of matrix questions should never need (or even be able) to stop in intermediate SpecCP where there is no relevant feature, and this allows them to simply skip the intervening SpecCP (wjhether it is locally [+wh] or not), thus deriving the well-known obviation of WH-islands in Bulgarian, shown in (22).

22) Bulgarian WH-island obviation:
   a.   **Kakvo** se čudiš   **koga** e   kupil Ivan ?
        what   wonder-2sg when   aux   bought Ivan
        "What do you wonder when Ivan bought?"
        (WH-island in English)

   b.   **Koga** se čudiš   **kakvo** e   kupil Ivan ?
        when   wonder-2sg what   aux   bought Ivan
        "When do you wonder what Ivan bought?
        (lower reading of *when* – WH-island in English)

In fact, we expect no Subjacency effects in Bulgarian at all with WH-movement, and it is possible that there are none (except for Complex NP constraint violations, which have a different character and constitute a much stronger violation in English as well). The account derives the WH island obviation effect as well. This is another welcome result.

   In the Russian/BCS/Polish case, since the Sidewards movement involves reintegration at the TP/Focus level, we don't expect WH-island obviation, and this is exactly what we find: the Russian equivalent in (23) are as bad as they are in English.

23) Russian WH-islands:

a. ***Čem** ty sprašivaeš', **kogda** Ivan zanimaetsja?
What-Instr you ask-2sg when Ivan studies
"What do you ask when Ivan studies?"

b. ***Kogda** ty sprašivaeš', **čem** Ivan zanimaetsja?
when you ask-2sg what Ivan studies
"When do you ask what Ivan studies? (lower reading of *when*)

We also might predict something else – since the WH-elements cluster at the TP edge, and if we assume they must do so in every clause (which is plausible given a certain view of FocP projections), we would expect them to behave as subjects (also a TP edge element) with regard to extraction over an overt complementizer. That is, we would expect a *that*-trace type effect with all embedded WH phrases in Russian and Polish, and this has in fact been observed, as shown in (24):

• Russian/Polish LD WH-movement out of indicatives:

24) a. ?? **Komu** ty dumaeš', čto Ivan pozvonil __ ?
who you think that Ivan called
"Who do you think that Ivan called?"

b. * **Kak** ty dumaeš', čto Ivan počinil mašinu ?
how you think that Ivan fixed car
"How do you think that Ivan fixed the car?"

If C is dropped, the effect disappears, as in (25):

25) a. **Komu** ty dumaeš', Ø Ivan pozvonil __ ?
who you think Ø Ivan called
"Who do you think that Ivan called?"

b. **Kak** ty dumaeš', Ø Ivan počinil mašinu ?
how you think Ø Ivan fixed car
"How do you think that Ivan fixed the car?"

This last application is fairly tentative, but seems promising, covering a restriction that is otherwise quite mysterious.

Such consequences for extraction and locality seem useful in helping us understanding how phrasal chunking takes place and when it is unsuccessful. At very least the account leads us to ask a different set of questions about locality effects and language types – we will expect multiple-movement languages to have different locality effects from single-movement languages and that appears to be what we find. Another possible advantage of the overall approach.

6   Conclusion

Under this account, there are two kinds of Superiority with different sources: English Superiority follows from Attract Closest, whereas Bulgarian Superiority follows from Bare Phrase Structure, the Extension Condition, and Sidewards Movement (all independently motivated). Happily, Tucking-In can be removed from the grammar. Superiority is non-parameterized and apparent lack of Superiority always results from independent factors.

More generally, we have a general theory of multiple movement as Self-motivated, requiring immediate satisfaction and hence Sidewards Movement (this piece, but only this one, follows Bošković 2007). Successive cyclicity cannot be forced in cases of Self-motivated Movement, resulting in some greater restrictions and some greater freedoms tha**n** Probe-driven movement, a rich area for further research.

References

Antonyuk, Svitlana 2015, Quantifier Scope and Scope Freezing in Russian. PhD Dissertation, Stony Brook University.

Bošković, Željko. 1997 "Superiority effects with multiple wh-fronting in Serbo-Croatian" *Lingua* 102: 1-20.

Bošković, Željko. 1998. Multiple wh-fronting and economy of derivation. *Proceedings of the West Coast Conference on Formal Linguistics* 16, 49-63. Stanford: Stanford University

Bošković, Zeljko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33(3):351-83.

Bošković, Zeljko. 2007. On the Locality and Motivation of Move and Agree: An Even More Minimal Theory. *Linguistic Inquiry* 38(4): 589-644.

Bruening, Benjamin. 2001. QR Obeys Superiority: Frozen Scope and ACD. *Linguistic Inquiry* 32 (2):233-73.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of Projection. *Lingua* 130: 33-49.

Grewendorf, Günter. 2001. Multiple Wh-Fronting. *Linguistic Inquiry* 32(1): 87-122.

Grewendorf, Günther, and Joachim Sabel. 1999. Scrambling in German and Japanese: Adjunction versus multiple specifiers. *Natural Language & Linguistic Theory* 17(1): 1-65.

Jaeger, T. Florian. 2004. Topicality and Superiority in Bulgarian wh-questions. In *Formal Approaches to Slavic Linguistics*, O. Arnaudova et al (eds.). Ann Arbor, Mich.: Michigan Slavic Publications.

Krapova, Iliyana, and Guglielmo Cinque. 2005. On the order of wh-phrases in Bulgarian multiple wh-fronting. *Working Papers in Linguistics, University of Venice* 15 171-195.

Lasnik, Howard, and Mamoru Saito. 1992. *Move–alpha*. MIT Press: Cambridge, MA.

Nunes, Jairo. 2001. Sidewards Movement. *Linguistic Inquiry* 31(2):303-344.

Nunes, Jairo. 2004. *Linearization of Chains and Sidewards Movement*. MIT Press.

Pesetsky, David. 2000. *Phrasal Movement and its Kin*. MIT Press.

Richards, Norvin 1997 *What moves where when in which languages?* PhD. Dissertation. MIT.

Richards, Norvin. 1998. The Principle of Minimal Compliance. *Linguistic Inquiry* 29(4): 599-629.

Richards, Norvin. 1999. Featural Specificity and the Ordering of Multiple Specifiers. In S. Epstein and N. Hornstein, eds. *Working Minimalism*, MIT Press: 127-158.

Rizzi, Luigi. 2004. Locality and Left Periphery. In A. Belletti (ed) *The Cartography of Syntactic Structures vol. 3*, Blackwell.

Rudin, Catherine. 1988. On Multiple Questions and Multiple WH-fronting. *Natural Language and Linguistic Theory* 6(4):445-501.

Scott, Tanya. 2012 *Whoever doesn't HOP must be Superior: the Left Periphery and the Emergence of Superiority in Russian* PhD Dissertation, Stony Brook University.

Simeonova, Vesela. 2012. Bulgarian multiple Wh-fronting: what happens at the left periphery? Paper given at the Alberta Conference on Linguistics – University of Lethbridge.

Stepanov, Artur. 1998. On wh-fronting in Russian. In NELS 28, GLSA, University of Massachusetts: 453-467.

Stepanov, Artur & Penka Stateva. 2009. When QR Disobeys Superiority. *Linguistic Inquiry* 40(1):176-85.

Stjepanović, Sandra. 1999. Multiple sluicing and superiority in Serbo-Croatian. *NELS 29*, 145–159.

john.bailyn@stonybrook.edu