Constraints on multiple specifiers

Andrew Murphy
andrew.murphy@uni-leipzig.de

Universität Leipzig

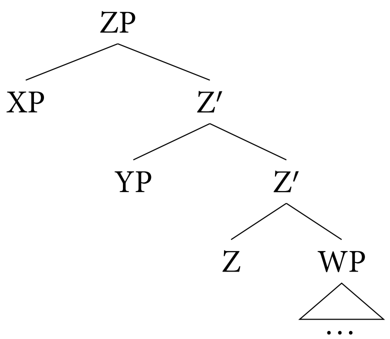
March 13, 2018

Abstract
This paper presents three examples of multiple fronting constructions in which creation of a second specifier is blocked for movement steps that also involve sub-extraction from NP. It is argued that these can be accounted for by assuming a violable constraint against multiple specifiers in the grammar. This constraint will be shown to interact with Left-Branch Extraction in Slavic, quantifier stranding in Korean and correlative fronting in Hindi to produce cumulative effects with multiple fronting. It will be demonstrated that these effects can be accounted for by extending the framework of Serial Harmonic Grammar to syntax. Furthermore, a strictly derivational approach to cumulative effects will be shown to successfully account for the observed asymmetries between subjects and objects.

1 Introduction

With the assumption of bare phrase structure, Chomsky (1995:245) proposed the abandonment of the so-called *Single Specifier Hypothesis* of X-bar Theory (Larson 1988:380f.; Speas 1990:79), stating that 'in principle, there might be a series of specifiers' (also see Koizumi 1995; Ura 1996; Mulders 1997; Zwart 1997; Nichols 1999; Doron & Heycock 1999; Chomsky 2000; Richards 2001; Rezac 2004; Lahne 2009). This meant that syntactic structures such as (1) became possible, where a head Z can project two specifiers containing XP and YP, respectively.

(1) *Multiple specifiers of a single head*:



Generally, it is assumed that multiple specifiers are a freely available option of UG, regularly created by intermediate movement via Spec-*v*P or object shift in Scandinavian (Chomsky 1995, 2000), for example. Furthermore, multiple specifiers of a single head have also be invoked as an explanation of constructions involving multiple displacement of elements within a single clause, for example with multiple wh-fronting in Bulgarian (e.g. Richards 2001) or multiple scrambling in Japanese (e.g. Grewendorf & Sabel 1999). In this paper, I discuss three distinct constructions from Serbo-Croatian, Korean and Hindi, in which an ordinarily available process of multiple fronting is blocked in conjunction with a particular kind of movement. Abstractly, all of these cases involve the configuration in (2). A second specifier of a head Z created by movement of XP is no longer licensed if it simultaneously involves sub-extraction from an NP.

(2) $^*[_{ZP}$ XP $[_{Z'}$ YP $[_{Z'}$ Z … [… $[_{NP}$ … $\_\__{XP}$ …] … $\_\__{YP}$ …]]]]

It will be argued the existence of a class of constructions exhibiting this restriction speaks in favour of a (violable) constraint in grammar that militates against the creation of a multiple specifiers of a single head. On a more descriptive level, the examples that will be discussed can be classified as cumulative effects, as defined in (3).

(3) *Cumulativity*:
A language allows process A and process B, but not the combination of A and B.

In the abstract case in (2), creation of multiple specifiers is generally possible, as is extraction from an NP, however not both simultaneously. An illustrative example of this involves the interaction between multiple wh-fronting and Left-Branch Extraction in Serbo-Croatian. As Section 2 will show, there are several Slavic languages that allow both multiple wh-fronting and Left-Branch Extraction independently, but not the co-occurence of these processes (4).

(4) a. *Čiji$_1$ kakva$_2$ $[_{NP}$ $\_\__1$ otac] kupuje $[_{NP}$ $\_\__2$ kola] ?
whose what.kind father buy car

b. *Kakva$_2$ čiji$_1$ $[_{NP}$ $\_\__1$ otac] kupuje $[_{NP}$ $\_\__2$ kola] ?
what.kind whose father buy car
'Whose father buys what kind of car?'

(*Serbo-Croatian*; Fernández-Salgueiro 2006:134)

Thus, this fits with the descriptive generalization in (3). This also follows the abstract pattern in (2); the second step of multiple fronting creating a multiple specifier of C also involves movement from inside an NP, regardless of the order of extraction:

(5) a. $^*[_{CP}$ wh_1 $[_{C'}$ wh_2 … $[_{vP}$ $[_{NP}$ $\_\__1$ NP] $[_{VP}$ V $[_{NP}$ $\_\__2$ NP]]]]]

b. $^*[_{CP}$ wh_2 $[_{C'}$ wh_1 … $[_{vP}$ $[_{NP}$ $\_\__1$ NP] $[_{VP}$ V $[_{NP}$ $\_\__2$ NP]]]]]

It will be shown that this makes an interesting prediction about mixed multiple fronting, i.e. with multiple fronting where only one movement step involves Left-Branch Extraction. If only subject extraction involves LBE, then this cannot be the second step of multiple fronting (6a). The reverse order of extraction is grammatical, however (6b). In other words, this leads to a

Superiority effect.

(6) a. *[$_{CP}$ *wh*$_1$ [$_{C'}$ *wh*$_2$ … [$_{vP}$ [$_{NP}$ __$_1$ NP] [$_{VP}$ V __$_2$]]]]

b. [$_{CP}$ *wh*$_2$ [$_{C'}$ *wh*$_1$ … [$_{vP}$ [$_{NP}$ __$_1$ NP] [$_{VP}$ V __$_2$]]]]

A strikingly similar pattern involving multiple scrambling can be found in Korean (Ko 2007, 2014). In general, scrambling of both a subject and an object is possible (7a), as is stranding of a subject quantifier (7b). However, multiple scrambling that involves quantifier stranding is blocked if it is the second step of fronting (7c), as predicted by (2). The reverse order of extraction is possible.

(7) a. [$_{CP}$ [$_{QP}$ S Q] [$_{C'}$ O … [$_{vP}$ __$_{QP}$ [$_{VP}$ __$_O$ V]]]]

b. [$_{CP}$ S … [$_{vP}$ [$_{QP}$ __$_S$ Q$_{SUB}$] [$_{VP}$ O V]]]

c. ?*[$_{CP}$ S [$_{C'}$ O … [$_{vP}$ [$_{QP}$ __$_S$ Q$_{SUB}$] [$_{VP}$ __$_O$ V]]]]

Furthermore, Ko (2007, 2014) reports that there is actually a subject/object asymmetry. Namely, unlike with subjects (7c), an object quantifier can be stranded as the second step of multiple scrambling (8).

(8) [$_{CP}$ O [$_{C'}$ S … [$_{vP}$ __$_S$ [$_{VP}$ [$_{QP}$ __$_O$ Q$_{OBJ}$] V]]]]

In fact, we will see that the same pattern also holds for multiple wh-fronting cases such as (6); LBE can be the second step of multiple fronting if it originates in object, rather than subject position. This will be shown to follow from a theory of cumulative constraint interaction which requires violations of constraints to occur local to the same movement step. Objects fundamentally differ from subjects in being merged inside the *v*P phase. For this reason, they have to undergo a step of intermediate movement to Spec-*v*P in order to be accessible for the higher phase. Crucially, the constraint violation for stranding can therefore be incurred at an intermediate, rather than final step and therefore fail to trigger a cumulative blocking effect. It will be shown that this asymmetry gives a principled account of the observed subject/object asymmetry with sub-extraction and multiple fronting.

The final example of the same phenomenon comes from correlative displacement in Hindi. (Bhatt 2003) shows that correlative clauses modifying a noun can undergo syntactic movement to be displaced from their associated demonstrative (9a). Furthermore, it is possible to have multiple correlative clauses in a single sentence (9b). However, Bhatt (2003) shows that multiple displacement of correlative clauses is not possible (9c), despite the fact that multiple scrambling of constituents is generally possible.

(9) a. $[_{\text{CP}}\ \text{CorCP}_1 \ldots [_{\text{NP}}\ \_\__{\text{CorCP}}\ \text{NP}_1\ ] \ldots]$

b. $[_{\text{CP}} \ldots [_{\text{NP}}\ \text{CorCP}_1\ \text{NP}_1\ ] \ldots [_{\text{NP}}\ \text{CorCP}_2\ \text{NP}_2\ ] \ldots]$

c. $^*[_{\text{CP}}\ \text{CorCP}_2\ [_{\text{C}'}\ \text{CorCP}_1 \ldots [_{\text{NP}}\ \_\__{\text{CorCP}_1}\ \text{NP}_1\ ] \ldots [_{\text{NP}}\ \_\__{\text{CorCP}_2}\ \text{NP}_2\ ] \ldots]]$

This restriction has the same explanation as the other two cases: both multiple scrambling and displacement of a correlative clause are possible individually, but not simultaneously. We can view this as a restriction on multiple specifier creation when it co-occurs with extraction of an NP adjunct (cf. Bhatt 2003:510). What is particularly interesting is the new observation that multiple specifiers of C created by intermediate movement do not trigger a cumulative effect. For example, multiple correlative displacement is possible if they do land in the same clause (10).

(10) $[_{\text{CP}}\ \text{CorCP}_2 \ldots [_{\text{CP}}\ \_\__{\text{CorCP}_2}\ [_{\text{C}'}\ \text{CorCP}_1 \ldots [_{\text{NP}}\ \_\__{\text{CorCP}_1}\ \text{NP}_1\ ] \ldots [_{\text{NP}}\ \_\__{\text{CorCP}_2}\ \text{NP}_2\ ]]]$

The reason for this effect (also found with successive-cyclic movement via *v*P in Serbo-Croatian and Korean) is that cumulative violation of constraints is strong enough to override the trigger for final, but not intermediate movement steps, which are driven by distinct constraints in the analysis that will be proposed.

In order to arrive at a unified account of these three phenomena as cumulative constraint violation triggered by multiple specifier creation, we require a theory of cumulativity. Although cumulative constraint interaction is sometimes proposed as an intuitive explanation (e.g. Chomsky 1973; Haegeman et al. 2014), there has no established theory of it for a derivational model of syntax. Here, I adopt a successful approach to cumulative effects from phonology, namely *(Serial) Harmonic Grammar* (e.g. Legendre et al. 1990; Pater 2009, 2016; Potts et al. 2010; Ryan 2017). In this optimality-theoretic framework, constraints bear weights, violations of which are deducted from a candidates harmony score. The general logic of the explanation is as follows: If a language allows a given process (e.g. Left-Branch Extraction, quantifier stranding or adjunct movement), then the respective constraint(s) against it (A) will have to bear a sufficiently low weight relative to the triggering constraint C in order to be licensed ($w(\text{C}) > w(\text{A})$). The same will hold for fronting of multiple constituents, which we assume violates a constraint B against creation of multiple specifiers of a head ($w(\text{C}) > w(\text{B})$). In the aforementioned cases in which these processes cannot be combined, the independently tolerable violations of the respective constraints 'gang up' to outweigh the penalty for non-application of a process ($w(\text{A+B}) > w(\text{C})$) and thereby block a particular movement step in the derivation. It will also been shown that a serial approach to optimization, where each stage of the derivation is evaluated, makes correct predictions about the distribution of cumulative effects, that is, that they only arise when both violations are incurred at the same derivational step. Furthermore, since intermediate and final steps of movement are driven by different constraints, we can account for the differences in permissible multiple specifier creation for these two movement types.

2 Multiple wh-fronting and Left-Branch Extraction in Slavic

The first phenomenon under consideration is the interaction of multiple wh-fronting and Left-Branch Extraction. While a subset of Slavic languages allow for both multiple wh-fronting and LBE, their combination is deemed ungrammatical. This is surprising under the view that each of these processes is a freely available option to the grammar. On the other hand, ascribing a tolerable, but nevertheless tangible, cost to each of these processes will allow us to treat the ban on multiple LBE as a cumulative effect.

2.1 Multiple Left-Branch Extraction

It is a well-known fact that there are languages which require what Kuno & Robinson (1972:478) call 'double dislocation' of wh-phrases to clause-initial position (Wachowicz 1974; Toman 1981; Comorovski 1986; Rudin 1988; Bošković 2002). Many Slavic languages exhibit this property, for example Serbo-Croatian (11a), Russian (11b) and Polish (11c).

(11) *Multiple wh-fronting*:

a. Ko$_1$ koga$_2$ __$_1$ vidi __$_2$?
who whom sees
'Who sees whom?' (*Serbo-Croatian*; Rudin 1988:449)

b. Kto$_1$ kogo$_2$ __$_1$ priglasil __$_2$ na užin ?
who who invited to dinner
'Who invited whom to dinner?' (*Russian*; Grebenyova 2012:21)

c. Kto$_1$ kogo$_2$ __$_1$ budzi __$_2$?
who whom wakes.up
'Who wakes up whom?' (*Polish*; Wachowicz 1974:158)

Furthermore, these languages also are among the sub-group of Slavic languages that allow so-called *Left-Branch Extraction* (Ross 1967; Corver 1990; Bošković 2005*b*) in which a prenominal modifier or possessor is sub-extracted from the noun phrase (12).

(12) *Left-Branch Extraction*

a. Čijeg$_1$ si vidio [$_{NP}$ __$_1$ oca] ?
whose are seen father
'Whose father did you see?' (*Serbo-Croatian*; Bošković 2005*a*:11)

b. Čju$_1$ on kupil [$_{NP}$ __$_1$ mašinu] ?
whose he bought car
'Whose car did he buy?' (*Russian*; Grebenyova 2012:83)

c. Czyjego$_1$ widziałeś [$_{NP}$ __$_1$ brata] ?
whose saw.2SG brother
'Whose brother did you see?' (*Polish*; Borsley 1983:340)

However, the combination of these two processes is not possible (see Fernández-Salgueiro 2006; Grebenyova 2012). In the Serbo-Croatian example in (13), multiple wh-movement of left-branches is not possible, regardless of the order of extraction.

(13) *No Multiple Left-Branch Extraction* (Serbo-Croatian; Fernández-Salgueiro 2006:134):

a. *Čiji$_1$ kakva$_2$ [$_{NP}$ —$_1$ otac] kupuje [$_{NP}$ —$_2$ kola] ?
whose what.kind father buy car

b. *Kakva$_2$ čiji$_1$ [$_{NP}$ —$_1$ otac] kupuje [$_{NP}$ —$_2$ kola] ?
what.kind whose father buy car
'Whose father buys what kind of car?'

Furthermore, other Slavic languages with multiple wh-fronting and Left-Branch Extraction, such as Russian (14) and Polish (15), also do not allow the combination of these two processes:

(14) *No Multiple Left-Branch Extraction* (Russian; Grebenyova 2012:82):

a. *Kakoj$_1$ čju$_2$ [$_{NP}$ —$_1$ aktër] kupil [$_{NP}$ —$_2$ mašinu] ?
which whose actor bought car

b. *Čju$_2$ kakoj$_1$ [$_{NP}$ —$_1$ aktër] kupil [$_{NP}$ —$_2$ mašinu] ?
whose which actor bought car
'Which actor bought whose car?'

(15) *No Multiple Left-Branch Extraction* (Polish):

a. *Czyja$_1$ którego$_2$ [$_{NP}$ —$_1$ matka] spotkała [$_{NP}$ —$_2$ studenta] ?
whose.NOM which.ACC mother.NOM met student.ACC

b. *Którego$_2$ czyja$_1$ [$_{NP}$ —$_1$ matka] spotkała [$_{NP}$ —$_2$ studenta] ?
which.ACC whose.NOM mother.NOM met student.ACC
'Whose mother met which student?'

This is particularly surprising for virtually all theories of wh-movement or Left-Branch Extraction. In most approaches, there is no relevant grammatical constraint against each of these individual processes and, as such, we would expect that they can combine freely. Descriptively, it will be argued that the ban on multiple LBE can be captured by the following constraint:

(16) *LBE Generalization*:
Left-Branch Extraction cannot be the second step of multiple wh-fronting.

This rules out multiple LBE because, regardless of the order of extraction, the second step will always necessarily violate (16).

There has not been much discussion of the ban on multiple LBE in previous literature. The only two analyses I am aware of are Fernández-Salgueiro (2006) and Grebenyova (2012). Both of these approaches have in common that they treat LBE as a fundamentally different type of movement from ordinary wh-movement. Grebenyova (2012) argues that the relevant difference is that LBE is head movement. In (17), the left-branch *kakuju* adjoins to the head of TopP.

(17) [$_{TopP}$ [$_{Top}$ Top+*kakuju*$_1$] … Ivan kupil [$_{NP}$ t$_1$ *knigu*]] ?
which Ivan bought book
'What kind of book did Ivan buy?' (Grebenyova 2012:87)

Subsequently, Grebenyova (2012:88) claims that 'multiple LBE is impossible due to [the] Head Movement Constraint'. As shown in (18), after the first left-branch has adjoined to Top$_1$, this now intervenes for movement to Top_2.

(18) *[$_{TopP_2}$ [$_{Top_2}$ Top_2+*kakuju*$_2$] [$_{TopP_1}$ [$_{Top_1}$ Top_1+*kakoj*$_1$] … [$_{NP}$ t_1 *student*] [$_{NP}$ t_2 *knigu*]]]
which whose student book

Appealing to the HMC in this way is problematic, however, since an ordinarily step of LBE-head movement such as (17) would have to be able to skip a number of intervening heads. Furthermore, this is is also an unusual conception of head movement, it is neither standard head-to-head movement (Travis 1984; Chomsky 1986), nor is it the kind of Head-to-Spec movement advocated by Vicente (2009) and Hein (2017), for example.

Fernández-Salgueiro (2006), on the other hand, assumes that LBE differs from ordinary wh-movement in the language in that it is driven by a [wh]-feature on C. He broadly follows the approach to mutliple wh-fronting in Bošković (1997*b*, 2002) and Stjepanović (1999) where wh-phrases undergo focus movement to their licensing position below CP (19).

(19) [$_{CP}$ C [$_{FocP}$ *wh*$_{[FOC]}$ *wh*$_{[FOC]}$ [… t_{wh} … t_{wh}]]]

This movement is Greed-based as it is driven by a focus feature on the wh-item itself, and not by some feature at the landing site. This is where he assumes (following Bošković 1997*a*:10f.) that LBE differs, arguing that movement of a left-branch is not driven by a focus feature on the moving item, but rather by an attracting [wh]-feature on the C head. The explanation for the incompatibility of multiple fronting and LBE is then rather simple: there is only a single wh-feature on the C head, so after the first left-branch has moved (20a), there is no trigger for the second movement step (20b).

(20) a. [$_{CP}$ *wh*$_1$ [$_{C'}$ $C_{[wh]}$ … [$_{NP}$ t_1 NP] … [$_{NP}$ wh_2 NP]]]

b. [$_{CP}$ *wh*$_2$ [$_{C'}$ *wh*$_1$ [$_{C'}$ $C_{[\cancel{wh}]}$ … [$_{NP}$ t_1 NP] … [$_{NP}$ t_2 NP]]]]

There are numerous challenges for this approach, however. For example, there are instances of focus LBE, meaning that it cannot always be driven by a [wh]-feature on C (see e.g. Bošković 2005*b*). Another problem with this theory is that we would expect left-branches to always precede non-left-branches in multiple extraction cases (21). This is because only the latter move to Spec-CP, the highest position.

(21) [$_{CP}$ *wh*$_1$ $C_{[\cancel{wh}]}$ [$_{FocP}$ *wh*$_2$ [… [$_{NP}$ t_1 NP] … t_2]]]

The data presented in the following section will show that, if anything, the reverse is true. What both of these approaches predict, due to their common assumption that LBE is simply a different type of movement, is that LBE should not interact with ordinary wh-movement. Since they are fundamentally different movement types (presumably driven by different features), there should be no interaction with regard to Minimality, for example. The following section will show that this prediction is not correct.

2.2 Superiority

The descriptive constraint in (16) makes a prediction with regard to what we might call 'mixed' multiple wh-fronting. This describes multiple wh-fronting in which only one of the movement steps involves LBE and has not been explicitly discussed in previous literature. Focusing on cases with an instance of subject LBE and ordinary object extraction such as (22), we see an interesting result. The only permissible order of multiple fronting is where the object precedes the subject left-branch (22a). The reverse order is ungrammatical (22b).[1]

(22) *Superiority with LBE from subject* (Serbo-Croatian):

a. (?)Šta$_2$ kakve$_1$ [$_{QP}$ dve [$_{NP}$ ___$_1$ devojke]] često čitaju ___$_2$?
what what.kind two girls often read

b. *Kakve$_1$ šta$_2$ [$_{QP}$ dve [$_{NP}$ ___$_1$ devojke]] često čitaju ___$_2$?
what.kind what two girls often read
'What do what kind of two girls often read?'

The same effect can also be seen with mixed multiple wh-fronting in Polish (23). Here, the remnant of subject LBE remains in postverbal position, thereby showing that the left-branch has moved. A similar pattern emerges: the object must precede the subject left-branch in order to be grammatical (23a).

(23) *Superiority with LBE from subject* (Polish):

a. (?)Kogo$_2$ czyja$_1$ spotkała [$_{NP}$ ___$_1$ matka] ___$_2$?
who.ACC whose.NOM met mother.NOM

b. *Czyja$_1$ kogo$_2$ spotkała [$_{NP}$ ___$_1$ matka] ___$_2$?
whose.NOM who.ACC met mother.NOM
'Whose mother met who?'

It is important to note that this is not due to subject LBE being somehow independently degraded (see Jurka 2010:187ff. for experimental evidence; also see Polinsky et al. 2013). Furthermore, we will see that mixed multiple fronting with object LBE behaves differently, for principled reasons, but for now we defer this discussion to Section 5. It seems that we are therefore dealing with what is essentially a Superiority effect. If a subject left-branch participates in mixed multiple fronting, then it must move first. This is particularly surprising since the languages in question are known

[1] These examples involve extraction across a quantifier to determine that the subject has indeed moved. Such cases of 'deep LBE' (Bošković 2005*b*) are normally ruled out, but quantifiers constitute an exception to this restriction (see e.g. Bošković 2012:205).

not to exhibit Superiority effects with clausemate extraction (24).

(24) a. Ko$_1$ koga$_2$ __$_1$ vidi __$_2$?
who whom sees

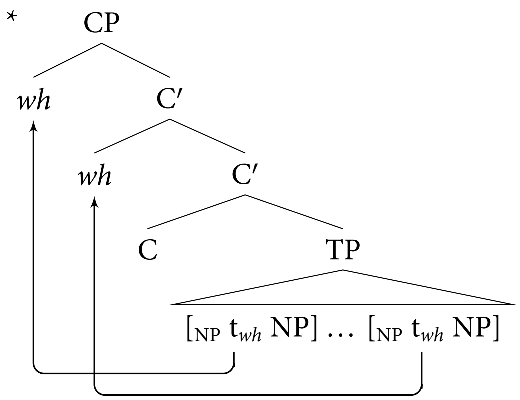
b. Koga$_2$ ko$_1$ __$_1$ vidi __$_2$?
whom who sees
'Who sees whom?' (*Serbo-Croatian*; Rudin 1988:473)

However, the emergence of what looks like a Superiority restriction follows naturally in light of the constraint in (16), repeated as (25). In Superiority-violating derivations such as (23b) with crossing dependencies, LBE is necessarily the second step of multiple wh-fronting, in contravention of (25).

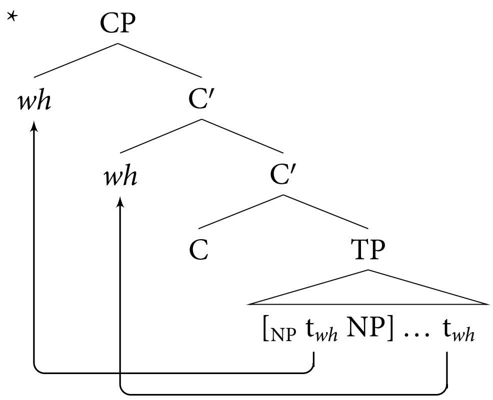
(25) *Left-Branch Extraction Generalization*:
Left-Branch Extraction cannot be the second step of multiple wh-fronting.

Abstractly, we see therefore see that the constraint in (25) can account for both the ban on multiple LBE and the emergence of Superiority with subject LBE. What they both have in common is that the second step of multiple fronting involves LBE. It will be shown that this can be viewed as a cumulative effect in which a second specifier of C cannot be created by a movement step involving LBE (26).

(26) *Multiple Left-Branch Extraction*:



Subject Left-Branch Extraction:

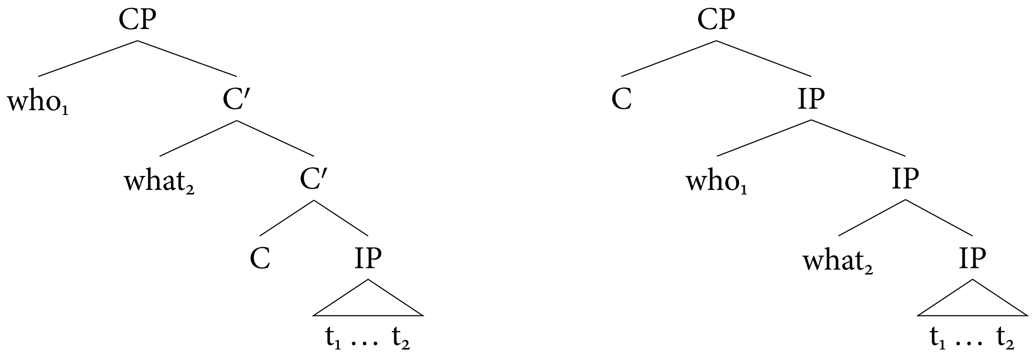


A direct consequence of this analysis, however, is that we have to assume that multiple wh-fronting in languages such as Serbo-Croatian involves movement to multiple specifiers of C. This is at odds with the traditional view of multiple fronting in Slavic going back to Rudin (1988). In the following section, I critically review some of the evidence for this position and show that there are no particularly strong arguments against the multiple specifier analysis of multiple wh-fronting in these languages.

2.3 On the syntax of multiple wh-fronting

Arguably, the standard view of multiple wh-fronting is that there exist two distinct structures. Some languages employ movement to multiple specifiers of Spec-CP (what Richards (2001) calls *CP-absorption* languages) whereas others adjoin some, or all, wh-phrases to a lower position (28).

(27) *CP-absorption*: *IP-absorption*:



This distinction goes back to the seminal paper by Rudin (1988), who proposed a [±MFS] parameter (=Multiply Filled SpecCP), with [+MFS] languages such as Bulgarian and Romanian moving wh-phrases to multiple specifiers of C and [−MFS] languages such as Serbo-Croatian and Polish, which lack this option. The evidence for this classification was based on the diagnostics in (28).

(28) *Typology of multiple wh-fronting languages* (Rudin 1988:478):

	[+MFS] Bulgarian, Romanian	[−MFS] Serbo-Croatian, Polish, Russian, etc.
multiple embedded extraction	✓	✗
wh-island violations	✓	✗
intervention by clitics, parentheticals, adverbs	✗	✓
Superiority effects	✓	✗

This view is incompatible with the account of the interaction of LBE with multiple fronting that was proposed in the previous section. However, there are a number of problems with the Rudin's classification, which undermine this basic distinction.

The first two diagnostics in (28) pertain to whether a language can license multiple specifiers of an embedded CP. Both the possibility to extract multiple phrases from an embedded clause and the permissibility of island violations necessitates the creation of a second specifier of CP.

(29) $[_{\text{CP}}\ \textit{wh} \ldots [_{\text{CP}}\ \textit{wh}\ [_{\text{C}'}\ \textit{wh}\ [_{\text{TP}}\ \text{t}_{\text{wh}} \ldots \text{t}_{\text{wh}}\ ]]]]$

The structure in (29) should be unavailable for [−MFS] languages, ruling out both multiple wh-fronting from embedded clauses and predicting sensitivity to wh-islands. The problem is that multiple wh-extraction from embedded clauses has been argued by Bošković (1997*a*, 2002, 2008*b*) to be possible for many speakers in Serbo-Croatian (this is actually explicitly acknowledged by Rudin 1988:453,fn.8).

(30) Ko$_1$ si koga$_2$ turdio [$_{CP}$ da je t$_1$ istukao t$_2$] ?
who 2SG whom claimed that is beaten
'Who did you claim beat whom?' (Bošković 1997*a*:5)

Furthermore, the acceptability of multiple embedded extraction has been reported for other [−MFS] languages such as Russian (Scott 2012) and Slovenian (Golden 1997) (again, this is subject to some speaker variability; see e.g. Mišmaš 2015). The other diagnostic relevant to the structure in (29) is that only [−MFS] should be sensitive to wh-islands. The problem here again is that some putative [−MFS] languages have been reported to permit wh-island violations (e.g. Polish; Cichocki 1983:64 and Czech; Rudin 1988:460). In support of this diagnostic, Rudin (1988:457) claims that Bulgarian, as a [+MFS] language, allows extraction from wh-islands, however this is actually rather restricted and only true for 'D-linked' wh-phrases and relative pronouns. In Bulgarian, movement of 'simplex' wh-phrases and adjuncts is robustly sensitive to to wh-islands (Rudin 1988:460; Bošković 2003:33), which is not predicted by the structure in (29). Even more problematically, Bošković (2003:34) points out that languages without multiple wh-fronting, such as Swedish, have exactly the same profile with regard to extraction from wh-islands as Bulgarian (Engdahl 1986; also see Bošković 2008*b*:262f. for similar data from a number of different languages). This suggests that the relevant factor must be linked to something other than the syntax of multiple wh-fronting. For this reason, we can disregard these diagnostics as evidence for the [±MFS] distinction.

The second set of diagnostics refer to the possibility for material such as clitics and parentheticals to intervene between fronted wh-phrases. In [+MFS] languages such as Bulgarian, clitics or parentheticals cannot intervene between wh-phrases in left-periphery (31a), whereas they can in [−MFS] languages such as Serbo-Croatian (31b).

(31) a. *Koj$_1$ *ti*$_2$ *e* kakvo$_3$ t$_1$ kazal t$_2$ t$_3$?
who 2SG.CL has what told
'Who told you what?' (*Bulgarian*; Rudin 1988:461)

b. Ko$_1$ *mu*$_2$ *je* šta$_3$ t$_1$ dao t$_2$ t$_3$?
who him 3SG.CL what given
'Who gave him what?' (*Serbo-Croatian*; Rudin 1988:462)

This is assumed to follow from respective structures for multiple fronting (32).

(32) a. [+MFS]: [$_{CP}$ [*wh wh wh*] [clitics/parentheticals [$_{TP}$ …]]]
b. [−MFS]: [$_{CP}$ (*wh*) [clitics/parentheticals [$_{TP}$ *wh wh* [$_{TP}$ …]]]]

However, there are also confounds in this domain. First, as Rudin (1988:462ff.) herself acknowledges, the two types of multiple-fronting languages also show independent differences in the type of clitics they have. In Bulgarian and Romanian, clitics always attach to the verb (also see Avgustinova 1994; Billings 2002; Franks 2008; Harizanov 2014). As such, the illicit placement of the clitics in (31a) does not support the structure in (32a) (also see Billings & Rudin 1996:54,fn.2). Furthermore, the obligatory placement of clitics between the wh-phrases in (31b) does not require the structure in (32b). Bošković (2001) argues at length that the placement of clitics in Serbo-Croatian is prosodically-driven, namely that clitics must occur in the second-position of an intonational phrase. Support for this comes from the fact that, if an optional pause is added between the fronted wh-phrases, then the clitic *je* can occur second in either of the intonational phrases (33a,b). Omission of the pause leads to placement after the first constituent (33c). In particular, the position of the clitic in (32b) is not predicted by the structure in (32b), but follows under the prosodic account.

(33) a. (Koji *je* čovjek$_1$)$_\iota$ (koju knjigu$_2$ t$_1$ kupio t$_2$)$_\iota$?
which 3SG.CL man which book bought

b. (Koji čovjek$_1$)$_\iota$ (koju *je* knjigu$_2$ t$_1$ kupio t$_2$)$_\iota$?
which man which 3SG.CL book bought

c. (Koji čovjek$_1$ *je* koju knjigu$_2$ t$_1$ kupio t$_2$)$_\iota$?
which man 3SG.CL which book bought
'Which man bought which book?' (Bošković 2001:70f.)

Thus, the example in (31b) is compatible with an analysis where wh-phrases move to Spec-CP and the clitic(s) are placed in second position at PF (34).

(34) $[_{\text{CP}}$ ko $[_{\text{C}'}$ koga $[_{\text{C}'}$ C $[_{\text{TP}} \ldots$ *je* $\ldots$]]]] $\Rightarrow$ PF: (((ko)$_\omega$ *je* (koga)$_\omega$ C)$_\varphi$ …)$_\iota$

A similar approach can be taken for intervention by parentheticals or adverbs between the wh-phrases, which is possible n Serbo-Croatian (35a), but is not in Bulgarian (35b) (Rudin 1988:468f.).

(35) a. Ko$_1$, po tebi, šta$_2$ t$_1$ pije t$_2$?
who by you what drinks
'Who, according to you, drinks what?' (*Serbo-Croatian*)

b. ?*Koj$_1$, spored tebe, kakvo$_2$ t$_1$ e kazal t$_2$?
who according to you what has said
'Who, in your opinion, said what?' (*Bulgarian*)

It is worth noting that this does not follow from the putative structural difference in (32). In order to account for the impenetrability of fronted wh-phrases in [+MFS] languages, an additional process of fusion or 'clustering' is necessary. While such a process is often assumed to apply before the wh-phrases reach their final landing site (e.g. Grewendorf 2001; Sabel 2001; Bailyn 2017), there is no principled reason why it could not apply after multiple fronting, as in (36) (cf. *m(orphological)-Merger*; Matushansky 2006 and *oblique movement*; Takano 2002:257). We can assume that, in [+MFS] languages, the fronted wh-phrases form a cluster via fusion in Spec-CP, which therefore prevents adjunction of a parenthetical at C′:[2]

(36) [$_{CP}$ who$_1$ [$_{C'}$ what$_2$ [$_{C'}$ C [$_{TP}$ t$_1$ … t$_1$]]]]
fusion

A [−MFS] language such as Serbo-Croatian would then simply lack this additional fusion process and therefore freely permit adjunction to C′. Some additional evidence in support of this view comes from the following contrast, which shows that Bulgarian does allow an intervening parenthetical if one of the fronted wh-phrases is complex (37b).

(37) a. **Koj*$_1$, spored tebe, *kakvo*$_2$ t$_1$ e kupil t$_2$?
who according to you what is bought
'Who, according to you, bought which book?'

b. ?*Koj*$_1$, spored tebe, *koja kniga*$_2$ t$_1$ e kupil t$_2$?
who according to you which book is bought
'Who, according to you, bought which book?' (Bošković 2002:361)

We can interpret this as the result of a constraint on the fusion operation in (36), namely that it can only apply to simplex elements. This is exactly the same restriction that Nunes (2004) proposes for fusion of a wh-phrase with a complementizer resulting in wh-copying in German, which is possible with simplex (38a), but not complex wh-phrases (38b).

(38) a. *Wen* glaubst du, *wen* sie liebt?
who believe you who she loves
'Who do you think she loves?'

b. **Welchen Mann* glaubst du, *welchen Mann* sie liebt?
which man believe you which man she loves
'Which man do you think she loves?'

(Fanselow & Mahajan 2000:220)

When fusion is blocked as in (37b), parentheticals are free to attach between the wh-phrases, as in [−MFS] languages.

The final of Rudin's (1988) diagnostics pertains to Superiority effects. In [+MFS] languages such as Bulgarian, a fronted subjects must precede a fronted object in the left periphery (39), while the order of fronted wh-phrases in [−MFS] languages such as Serbo-Croatian is free (24).

(39) a. Koj$_1$ kogo$_2$ t$_1$ vižda t$_2$?
who whom sees

b. *Kogo$_2$ koj$_1$ t$_1$ vižda t$_2$?
whom who sees
'Who sees whom?' (*Bulgarian*; Rudin 1988:472f.)

[2] Independent evidence for adjunction of parentheticals to C′ can be seen in the following examples from English where a parenthetical intervenes between a fronted phrase in Spec-CP and an auxiliary in C:

(i) a. Who, in your opinion, did Mary suspect?
b. Never, in my opinion, was so much owed by so many. (Wilder 1997:331)

The crucial difference is that wh-movement in Bulgarian-type languages involves Ā-movement to multiple specifiers of a single C head, whereas in Serbo-Croatian-type languages, matrix wh-movement is actually adjunction to TP. Furthermore, he argues that Superiority effects follow as a direct result of this structural difference. In particular, Richards (1999, 2001) proposed the concept of *tucking-in*, whereby movement to the second specifier of a head creates a specifier below the first. Since Bulgarian is a CP-absorption language, multiple wh-movement will always result in tucking-in, and therefore order preservation. Serbo-Croatian, on the other hand, is an IP-absorption language where movement is scrambling-like, adjunction movement and each movement dependency is entirely independent of the others, leading to a general lack of Superiority with multiple wh-fronting. The main problem, though, is that this correlation is not perfect. We do in fact find Superiority effects in a variety of places in [−MFS] languages. For example, Bošković (1997*a*, 2002) and Stjepanović (2003) show that there are several contexts in which Superiority effects can be found in Serbo-Croatian, i.e. with long-distance multiple extraction, embedded questions, multiple sluicing, correlatives, *li*-constructions and in clauses with topicalized constituents (also see Scott 2012 for similar claims for Russian). One example of Superiority in Serbo-Croatian is given in (40) with an overt interrogative C head, which is realized as the second-position clitic *li*.

(40) a. Ko *li* koga t_1 voli t_2 ?
who C whom loves

b. *Koga *li* ko t_1 voli t_2 ?
whom C who loves
'Who on earth loves whom?' (Bošković 2002:354)

To this list, we can also add the emergence of Superiority with subject LBE observed in (22). The approach suggested by Bošković (2002:354) for contexts showing exceptional Superiority in [−MFS] languages is to treat this as an idiosyncratic property of C. In other words, some C heads bear a [wh]-probe feature that leads to attraction of the closest wh-phrase. In constructions showing no Superiority restrictions, i.e. clausemate extraction in matrix clauses, wh-phrases undergo 'Greed-based' focus movement triggered by a focus feature on the wh-items themselves (Bošković 1999:167). Since these movement steps are independent instances of adjunction to TP, they can apply in either order (41).

(41) a. $[_{CP}\ \mathrm{C}\ [_{TP}\ wh_{2[\mathrm{FOC}]}\ [_{TP}\ wh_{1[\mathrm{FOC}]}\ [_{TP}\ \dots\ t_1\ \dots\ t_2\ ]]]]$

b. $[_{CP}\ \mathrm{C}\ [_{TP}\ wh_{1[\mathrm{FOC}]}\ [_{TP}\ wh_{2[\mathrm{FOC}]}\ [_{TP}\ \dots\ t_1\ \dots\ t_2\ ]]]]$

This becomes problematic in light of the generalization in (16), repeated as (42), which we saw can derive both the ban on multiple LBE and the emergence of Superiority with subject LBE.

(42) *LBE Generalization*:
Left-Branch Extraction cannot be the second step of multiple wh-fronting.

In an analysis such as (41) where wh-phrases undergo independent steps of Greed-based movement, the 'second step of multiple wh-fronting' does not have any clear status. Thus, implementing (42) as a strictly local constraint that only makes reference to properties of a single movement step becomes incredibly difficult. In a theory where multiple fronting involves cyclic movement to multiple specifiers, however, the second step of multiple wh-fronting always has the inherent property of creating an additional specifier of interrogative C. It is this fact that the cumulative analysis to follow will try to exploit.

The conclusion of the preceding discussion is that the evidence for a distinction between two types of multiple wh-fronting languages is actually rather weak. Many of the diagnostics turn out to not be relevant to the proposed syntactic difference, and the others have alternative, perhaps even preferable, explanations. The consequence of this is that a unified account of multiple wh-fronting in Slavic, where wh-phrases move to multiple specifiers of C, is possible. In order to account for Superiority, we must take the more nuanced view that it is some inherent trait of the C head that gives rise to order preserving movement.[3] In [+MFS] languages, this holds for all instances of interrogative C, whereas this is a construction-specific property of certain C heads in [−MFS] languages such as overt *li* in Serbo-Croatian (40). Assuming that putative [−MFS] languages like Serbo-Croatian and Polish also have movement to multiple specifiers of C allows us to formalize the descriptive generalization in (42) as the derivational constraint that Left-Branch Extraction cannot create a second specifier of C. This can be analyzed as a cumulative effect in that a movement step may either involve LBE or creation of a multiple specifier, but not both simultaneously. In order to achieve a formal analysis of this, we first require an explicit theory of cumulative constraint interaction. This is what the following section will provide.

3 A theory of cumulativity

In order to provide an analysis of restrictions on multiple specifier creation in terms of cumulative effects, we first require an explicit theory of cumulativity. While the notion of cumulative constraint interaction has been proposed at various points in the literature (see e.g. Chomsky 1973:239,fn.19; Ross 1987; Haegeman et al. 2014), previously proposed theories primarily focus on deriving gradience in acceptability judgements rather than blocking a given derivational step, as required for the puzzles at hand (e.g. Keller 2000, 2006; Jäger & Rosenbach 2006; Adli 2011). I will propose that what has been a successful framework for analyzing cumulative blocking effects in phonology, namely *Serial Harmonic Grammar* (e.g. Kimper 2011, 2016; Pater 2012; Kaplan 2016; Ryan 2017), can equally account for cumulativity in syntax. This framework consists of two major components: weighted constraints from Harmonic Grammar and serial optimization from Harmonic Serialism. I will present each in turn in the following sections.

[3] This could still be implemented in terms of a counter-cyclic 'tucking in' derivation, or by means of some other theory of order preservation such as *Shape Conservation* (Williams 2003; also see Müller 2001), for example.

3.1 Harmonic Grammar

At the core of optimality-theoretic approaches is the assumption of violable constraints (Prince & Smolensky 1993/2004). The fundamental idea is that, among a set of competing candidates, the optimal output is determined based on the evaluation of their relative harmony based on a set of ranked, violable constraints. In an alternative predecessor to OT, *Harmonic Grammar* (HG) (Legendre et al. 1990, 2006; Pater 2009, 2014, 2016; Potts et al. 2010; Jesney 2016), constraints are not ranked, but instead bear weights. These weights are deducted from a candidates harmony score, and the candidate with the highest harmony score is selected as optimal. To see this, consider the following basic syntactic example involving wh-movement. Driving wh-movement, we have the markedness constraint Wh-Crit(erion) (43a) that requires wh-phrases to be in Spec-CP. The counteracting constraint Stay (43b) penalizes movement.

(43) a. Wh-Criterion (Rizzi 1996):
Wh-phrases must be in the specifier of a licensing head $C_{[wh]}$.
b. Stay (Ackema & Neeleman 1998:448):
Do not move (i.e. assign a violation for each trace/copy).

In a language with wh-movement, Wh-Crit must outrank Stay. In HG terms, the weight of Wh-Crit must be higher than that of Stay. As (44) shows, the penalty incurred for applying wh-movement (44a) is worse (−3) than the cost of a violation of Stay (44b) (−2). For this reason, wh-movement is licensed.

(44)

$[_{CP}\ C_{[wh]} \dots [_{VP}\ V\ wh_1\ ]]$	Wh-Crit $w=3$	Stay $w=2$	$\mathcal{H}$
a. $[_{CP}\ C_{[wh]} \dots [_{VP}\ V\ wh_1\ ]]$	−1		−3
☞ b. $[_{CP}\ wh_1\ [_{C'}\ C_{[wh]} \dots [_{VP}\ V\ t_1\ ]]]$		−1	−2

One core way in which HG differs from Standard OT is that there is no *strict domination*. This means that violations of less important constraints can ‘gang up’ to outweigh a violation of some more important constraint. It is this property of HG that gives a natural explanation of cumulative effects. Recall the definition of *cumulativity* in (3), repeated below.

(45) *Cumulativity*:
A language allows process A and process B, but not the combination of A and B.

As we saw above for wh-movement, a legitimate grammatical process comes at the expense of a violation of some less important constraint relative to the constraint triggering the operation. For process A, let us assume the trigger constraint C bears a higher weight (w=3) than the constraint *A violated by application of A (w=2). This means that non-application of process A comes at a higher cost than the candidate carrying out A (46).

(46) *Process A permitted*:

Input	C $w = 3$	*A $w = 2$	$\mathcal{H}$
a. Process A doesn't apply	−1		−3
☞ b. Process A applies		−1	−2

If there is another process B that is also driven by C, we can assign the same weight to the corresponding constraint *B (47).

(47) *Process B permitted*:

Input	C $w = 3$	*B $w = 2$	$\mathcal{H}$
a. Process B doesn't apply	−1		−3
☞ b. Process B applies		−1	−2

Given the current set of weights, both violations of *A and *B are tolerable individually, but if the processes A and B co-occur, then the combined sum of their violations (−4) results in a worse harmony score than violating C (the trigger for the operations A and B) (48), i.e. −3.

(48) *Co-occurence of A and B prohibited (gang effect)*:

Input	C $w = 3$	*A $w = 2$	*B $w = 2$	$\mathcal{H}$
☞ a. Process A/B doesn't apply	−1			−3
b. Process A/B applies		−1	−1	−4

In this way, we can derive basic signature of cumulativity in (45). This is also known as a *gang effect* since two less important constraints that would not be able to affect the outcome individually cooperate or 'gang up' to block the application of a process driven by a higher-ranked constraint.[4]

[4] At this point, it is important to mention that the actual values we choose as the weight for a given constraint is arbitrary. What instead matters is that particular *weighting conditions* hold between constraints. In order to have a gang effect as in (48), the weights constraints *A and *B must be individually lower than C (ia,b), but not their sum (ic).

(i) a. $w(\text{C}) > w(\text{*A})$
 b. $w(\text{C}) > w(\text{*B})$
 c. $w(w(\text{*A})+w(\text{*B})) > w(\text{C})$

For present purposes, any set of weights compatible with (i) can be chosen.

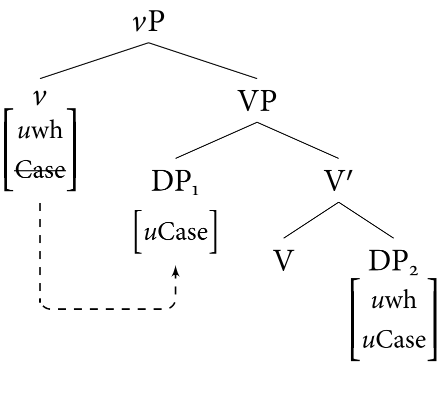
3.2 Harmonic Serialism

The second important property of Serial Harmonic Grammar is cyclic optimization. This refers to the assumption that only a single change can be made to the input at a time. The winner of a given optimization is then subject to iterative subsequent optimizations until no further improvements are possible. This framework is known as *Harmonic Serialism* (McCarthy 2000, 2008*a*,*b*, 2010, 2016; Torres-Tamarit 2012; Elfner 2016). The result is a derivational theory where each continuation of a derivation is determined by ranked or weighted constraints. Arguably, this is what a standard derivational approach to Minimalist syntax, such as that in Chomsky (1995, 2000, 2001) requires. Applications of HS to syntax have been shown by Heck & Müller (2003, 2013, 2016) to have several welcome consequences. In particular, cyclic optimization gives an explicit theory for determining possible continuations of a given derivational step. This is required for implementing what Müller & Sternefeld (2001:8) call *translocal* economy, i.e. competition between possible (intermediate) output representations (i.e. reference-set economy; Chomsky 1995:227). To see an example of this, consider the economy constraint *Multitasking* in (49).

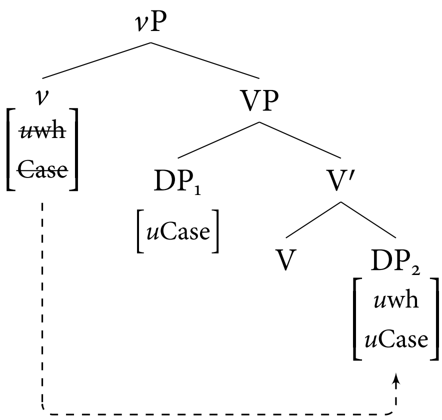
(49) *Multitasking* (van Urk & Richards 2015:132; Richards 2016:135):
If two Agree operations A and B are possible, and the features checked by A are a superset of those checked by B, the grammar prefers A.

In their analysis of Dinka, van Urk & Richards (2015:131f.) appeal to this constraint to resolve the following abstract derivational scenario. They assume that *v* bears both a wh-probe and a Case probe [*u*wh, Case]. Now, imagine that either argument of a ditransitive is a potential goal for Agree. One option is to agree with the indirect object that bears a corresponding Case feature (50) (the option favoured by pure Minimality considerations). The [*u*wh] would then be checked on the second cycle of Agree with the direct object. However, agreeing with the direct object straight away checks both features simultaneously (51). Given (49), this latter option is preferable, despite violating strict Minimality.

(50) *Option 1*:



(51) *Option 2*:



Translating the *Multitasking* constraint into Harmonic Serialism is actually rather trivial. The desire to Agree with the goal that checks the most features follows from the basic desire to maximize constraint satisfaction. We can assume that agreement operations are driven by a constraint

AGREE that is violated by each unchecked probe feature on a head. As long as AGREE bears a higher weight than the constraint against Minimality (e.g. the MINIMALLINKCONDITION), then the candidate avoiding violations of AGREE will be favoured. In (52), we have reached the derivational step where *v* is merged and have the option of doing nothing (52a) (the faithful candidate), agreeing with IO (52b) or agreeing with the DO (52c). Given the preference expressed by the weighting conditions in (52), the option that maximizes the number of checked features (52c) is preferred, even if this violates Minimality.

(52) *Multitasking in Serial Harmonic Grammar*:

$[_{vP}\ v_{[*\text{Case},u\text{wh}*]}\ [_{\text{VP}}\ \text{DP}_{[u\text{Case}]}\ [_{\text{V}'}\ \text{V}\ \text{DP}_{[\text{wh},u\text{Case}]}\ ]]]$	AGREE $w = 3$	MLC $w = 2$	$\mathcal{H}$
a. $[_{vP}\ v_{[*\text{Case},u\text{wh}*]}\ [_{\text{VP}}\ \text{DP}_{[u\text{Case}]}\ [_{\text{V}'}\ \text{V}\ \text{DP}_{[\text{wh},u\text{Case}]}\ ]]]$	−2		−6
b. $[_{vP}\ v_{[*u\text{wh}*]}\ [_{\text{VP}}\ \text{DP}_{[u\text{Case}]}\ [_{\text{V}'}\ \text{V}\ \text{DP}_{[\text{wh},u\text{Case}]}\ ]]]$	−1		−3
☞ c. $[_{vP}\ v\ [_{\text{VP}}\ \text{DP}_{[u\text{Case}]}\ [_{\text{V}'}\ \text{V}\ \text{DP}_{[\text{wh},u\text{Case}]}\ ]]]$		−1	−2

Taken together, these two assumptions provide a suitable theoretical framework for analyzing the blocking of multiple specifier creation as gang effects triggered at a particular derivational step. The following will section discusses the exact nature of the constraints involved.

4 Multiple wh-fronting and LBE

4.1 Ruling out Multiple Left-Branch Extraction

Recall that, although languages such as a Serbo-Croatian have both multiple wh-fronting and LBE individually, the combination of these processes is not possible (53).

(53) *No Multiple Left-Branch Extraction* (Serbo-Croatian):

a. *Kakva$_2$ čiji$_1$ [$_{\text{NP}}$ ——$_1$ otac] kupuje [$_{\text{NP}}$ ——$_2$ kola] ?
what.kind whose father buy car

b. *Čiji$_1$ kakva$_2$ [$_{\text{NP}}$ ——$_1$ otac] kupuje [$_{\text{NP}}$ ——$_2$ kola] ?
whose what.kind father buy car
'Whose father buys what kind of car?'

With the theory outlined in the preceding section, we can now treat this as a gang effect. The guiding idea is that, in the relevant languages, both multiple wh-fronting and LBE come at the expense of a tolerable violation of a constraint. While the violations of these constraints may be incurred individually, simultaneous violation becomes too costly. In a language allowing LBE, the constraint responsible for driving wh-movement WH-CRITERION that we saw in (43a) bears a higher weight than the constraint against extraction of left-branch modifiers. We will simply call this constraint LeftBranchCondition (Lbc).

(54) LeftBranchCondition (cf. Ross 1986:127f.):
Assign a violation for a syntactic object γ in position α in $[\ \alpha \dots\ [_{\text{NP}}\ [_{\text{N}'}\ \text{N} \dots\ ]]\ \dots\ ]$, where γ corresponds to β in the input $[\ \dots\ [_{\text{NP}}\ \beta\ [_{\text{N}'}\ \text{N} \dots\ ]]\ \dots\ ]$.

Note that this constraint implements Ross' *Left-Branch Condition* as a faithfulness constraint against movement of items from the specifier position of NP. Other approaches that try to derive LBE from the lack of a DP phase (e.g. Bošković 2005*b*) assume that LBE comes for free in NP languages and therefore make it difficult to implement a cumulative analysis.[5] If we give Wh-Crit and Lbc weights of 3 and 2 respectively, then this grammar will permit LBE (55b).

(55) *Left-Branch Extraction possible*:

$[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_1\ \text{NP}]\ ]]$	Wh-Crit $w = 3$	Lbc $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_1\ \text{NP}]\ ]]$	−1		−3
☞ b. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ ]]]$		−1	−2

Assuming, as motivated in Section 2.3, that multiple wh-fronting targets multiple specifiers of C, multiple wh-fronting will violate the following markedness constraint against multiple specifiers of the same head:

(56) *Mult(iple)-Spec(ifier):
Multiple specifiers of a single head are prohibited.
$^*[_{\text{XP}}\ \alpha\ [_{\text{X}'}\ \beta\ [_{\text{X}'}\ \text{X}\ [_{\text{YP}} \dots\ ]]]]$

Again, as long as this constraint bears a lower weight than Wh-Crit, then multiple wh-fronting will be licensed (57b). In a derivational approach, the first step of multiple fronting removes one of the violations of Wh-Crit by moving a wh-phrase in Spec-CP.

(57) *Multiple wh-fronting possible* (Step 1):

$[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ wh_1 \dots wh_2\ ]]$	Wh-Crit $w = 3$	*MultSpec $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ wh_1 \dots wh_2\ ]]$	−2		−6
☞ b. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ \text{t}_1 \dots wh_2\ ]]]$	−1		−3

Given the assumption of cyclic optimization, the optimal output in (57b) is evaluated once more. Here, the second step of multiple wh-fronting creates a second specifier of C at the cost of a

[5]The caveat here is that NPs do have phasal status, namely that only the highest edge is accessible (Bošković 2016*a*). It is actually the lack of intermediate movement to DP that leads to the legitimacy of LBE. Nevertheless, there is still no constraint in the grammar that penalizes extraction from outer edges in such an approach.

tolerable violation of *Mult-Spec (58b).

(58) *Multiple wh-fronting possible* (Step 2):

$[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ t_1 \dots wh_2\ ]]]$	Wh-Crit $w = 3$	*MultSpec $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ t_1 \dots wh_2\ ]]]$	−1		−3
☞ b. $[_{\text{CP}}\ wh_2\ [_{\text{C}'}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ t_1 \dots t_2\ ]]]]$		−1	−2

Given the condition that the summed weights of Lbc and *Mult-Spec are higher than Wh-Crit, we can restate the descriptive generalization in (45) in more technical terms, as in (59).

(59) *LBE Generalization (revised)*:
A single step of wh-movement cannot violate both *Mult-Spec and Lbc simultaneously.

To see how this rules out multiple LBE, consider the following cyclic derivation. At the point of the derivation in which interrogative C is merged, we have the option to move one of the wh-phrases. For present purposes, it does not matter which one we move first. In (60b), movement of the left-branch wh_1 trades a violation of Wh-Crit against a less costly violation of Lbc and is therefore the optimal output.

(60) *Multiple Left-Branch Extraction* (Step 1):

$[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ wh_1\ \text{NP}] \dots [_{\text{NP}}\ wh_2\ \text{NP}]\ ]]$	Wh-Crit $w = 3$	Lbc $w = 2$	*MultSpec $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ wh_1\ \text{NP}] \dots [_{\text{NP}}\ wh_2\ \text{NP}]\ ]]$	−2			−6
☞ b. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ t_1\ \text{NP}] \dots [_{\text{NP}}\ wh_2\ \text{NP}]\ ]]]$	−1	−1		−5

This output forms the input to the subsequent optimization in (61). Here, we now try to move the second left-branch wh_2 to fully satisfy Wh-Crit. However, this movement step in (61b) now violates both Lbc, due to it being movement of a left-branch, and *Mult-Spec as the second step of multiple fronting. It therefore triggers a gang effect, since the summed violations of the two constraints lead to a worse harmony score than simply not moving at all (61a).

(61) *Multiple Left-Branch Extraction* (Step 2):

$[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ t_1\ \text{NP}] \dots [_{\text{NP}}\ wh_2\ \text{NP}]\ ]]]$	Wh-Crit $w = 3$	Lbc $w = 2$	*MultSpec $w = 2$	$\mathcal{H}$
☞ a. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ t_1\ \text{NP}] \dots [_{\text{NP}}\ wh_2\ \text{NP}]\ ]]]$	−1			−3
b. $[_{\text{CP}}\ wh_2\ [_{\text{C}'}\ wh_1\ [_{\text{C}'}\ \text{C}_{[\text{wh}]} \dots [_{v\text{P}}\ [_{\text{NP}}\ t_1\ \text{NP}] \dots [_{\text{NP}}\ t_2\ \text{NP}]\ ]]]]$		−1	−1	−4

For this reason this movement step of LBE is not licensed and we correctly rule out multiple LBE. It should be clear that the reverse order of extraction would have led to the same result, since the second step of multiple fronting will inevitable be LBE. Selecting the optimal output in (61a) will ultimately lead to a crash at the interfaces, as discussed further in Section 4.4.

4.2 The emergence of Superiority

With cases of mixed multiple wh-fronting, we saw that things were different. Recall from (22), repeated below, that subject LBE and ordinary object wh-movement we find a Superiority effect, the only legitimate derivation is one in which the subject left-branch moves first.

(62) *Superiority with LBE from subject* (Serbo-Croatian):

a. (?)Šta$_2$ kakve$_1$ [$_{QP}$ dve [$_{NP}$ __$_1$ devojke]] često čitaju __$_2$?
what what.kind two girls often read

b. *Kakve$_1$ šta$_2$ [$_{QP}$ dve [$_{NP}$ __$_1$ devojke]] često čitaju __$_2$?
what.kind what two girls often read
'What do what kind of two girls often read?'

This restriction can also be derived by the same grammar that rules out multiple LBE, since it is based on the same descriptive generalization that the second of multiple wh-fronting cannot involve LBE, as (62b) does. Consider first the Superiority-respecting derivation in which the subject left-branch moves first. In the first step, a violation of Wh-Crit is traded against a violation of Lbc (63b), constituting harmonic improvement. The subsequent step involves movement of the wh-object wh_2, which only violates *Mult-Spec due to creation of a second specifier of C.

(63) *Superiority-respecting derivation* (Step 1):

[$_{CP}$ … [$_{vP}$ [$_{NP}$ wh_1 NP] [$_{v'}$ v [$_{VP}$ V wh_2]]]]]	Wh-Crit $w = 3$	*Mult-Spec $w = 2$	Lbc $w = 2$	$\mathcal{H}$
a. [$_{CP}$ … [$_{vP}$ [$_{NP}$ wh_1 NP] [$_{v'}$ v [$_{VP}$ V wh_2]]]]]	−2			−6
☞ b. [$_{CP}$ wh_1 … [$_{vP}$ [$_{NP}$ t$_1$ NP] [$_{v'}$ v [$_{VP}$ V wh_2]]]]]	−1		−1	−5

Superiority-respecting derivation (Step 2):

[$_{CP}$ wh_1 … [$_{vP}$ [$_{NP}$ t$_1$ NP] [$_{v'}$ v [$_{VP}$ V wh_2]]]]	Wh-Crit $w = 3$	*Mult-Spec $w = 2$	Lbc $w = 2$	$\mathcal{H}$
a. [$_{CP}$ wh_1 … [$_{vP}$ [$_{NP}$ t$_1$ NP] [$_{v'}$ v [$_{VP}$ V wh_2]]]]	−1			−3
☞ b. [$_{CP}$ wh_2 [$_{C'}$ wh_1 … [$_{vP}$ [$_{NP}$ t$_1$ NP] [$_{v'}$ v [$_{VP}$ V t$_2$]]]]]		−1		−2

This derivation is licit because no single step of the derivation violates both Lbc and *Mult-Spec, as prohibited by (59). Things are different with the Superiority-violating derivation in (64), where the non-closest wh-phrase moves first. The first step of extraction of the wh-object wh_2 does not violate either *Mult-Spec or Lbc. However, the second step now violates both constraints and triggers a gang effect, as with multiple LBE.

(64) *Superiority-violating derivation* (Step 1):

$[_{\text{CP}} \ldots [_{v\text{P}} [_{\text{DP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_2\ ]]]]$	Wh-Crit $w = 3$	*Mult-Spec $w = 2$	Lbc $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}} \ldots [_{v\text{P}} [_{\text{DP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_2\ ]]]]$	−2			−6
☞ b. $[_{\text{CP}}\ wh_2 \ldots [_{v\text{P}} [_{\text{NP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]$	−1			−3

Superiority-violating derivation (Step 2):

$[_{\text{CP}}\ wh_2 \ldots [_{v\text{P}} [_{\text{NP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]$	Wh-Crit $w = 3$	*Mult-Spec $w = 2$	Lbc $w = 2$	$\mathcal{H}$
☞ a. $[_{\text{CP}}\ wh_2 \ldots [_{v\text{P}} [_{\text{NP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]$	−1			−3
b. $[_{\text{CP}}\ wh_1\ [_{\text{C}'}\ wh_2 \ldots [_{v\text{P}} [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]]$		−1	−1	−4

This shows that how both the ban on multiple LBE and Superiority with subject LBE can be accounted for by the same analysis. In the preceding discussion, we did not include intermediate movement of the wh-object to *v*P. This is not an oversight, however, and will in fact play a crucial role in deriving the subject/object asymmetry to be discussed in Section 5.

4.3 Serialism vs. parallelism

In the preceding analysis, the crucial generalization is that a single step of multiple wh-movement cannot violate both *Mult-Spec and Lbc. We saw that this follows naturally in a serial approach to optimization, where evaluation applies at each derivational step. It is particularly interesting to note that the result is different if we translate the analysis into a parallel account, where all movement steps apply simultaneously. As shown in (65), Parallel HG predicts no difference with regard to the order of extraction with subject LBE.

(65) *Wrong prediction of Parallel HG*:

$[_{\text{CP}} \ldots [_{v\text{P}} [_{\text{DP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_2\ ]]]]$	Wh-Crit $w = 3$	Lbc $w = 2$	*MultSpec $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}} \ldots [_{v\text{P}} [_{\text{NP}}\ wh_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_2\ ]]]]$	−2			−6
☞ b. $[_{\text{CP}}\ wh_2\ wh_1 \ldots [_{v\text{P}} [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]$		−1	−1	−4
☞ c. $[_{\text{CP}}\ wh_1\ wh_2 \ldots [_{v\text{P}} [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]$		−1	−1	−4

This is because a parallel optimization can identify the locus of each of violation in order to identify whether both violations are incurred by the same movement step. Since the generalization about LBE is inherently derivational, it can be straightforwardly expressed in a serial, but not a parallel approach, without enriching representations or constraint definitions significantly. The analysis developed here therefore shows the potential advantage of a derivational, as opposed to a parallel architecture of grammar.

4.4 Intermediate movement steps and ineffability

So far, we have not talked about intermediate movement of the object to Spec-*v*P, as is standardly assumed in phase theory (Chomsky 2000, 2001). One issue about this kind of successive-cyclic movement is how to account for the Look Ahead problem associated with it. In other words, we have to know at a sufficiently early stage of the derivation that a particular probe will be merged at some later point in order to trigger movement to the phase edge. Heck & Müller (2003) address this problem in Harmonic Serialism by proposing the following constraint:

(66) PhaseBalance (Heck & Müller 2003:105):
For every feature [∗F∗] in the numeration, there must be an accessible feature [F] at the phase level (i.e. at the phase edge) or in the numeration.

The central idea is that intermediate movement is a repair to the constraint in (66). The input to a given derivation also contains the numeration of remaining elements to be merged and so it can be locally determined whether or not there is a potential checker at the phase edge. In the input to (67), there is no potential wh-feature at the edge of *v*P as a checker for the [∗wh∗] probe feature on C. This means that (67a) violates PhaseBalance. As a repair, the wh-phrase is moved to the edge of *v*P (67b).

(67) *Intermediate movement of wh-object*:

$[_{vP}$ DP $[_{v'}$ v $[_{VP}$ V wh_1]]] ⊕ {T, $C_{[*wh*]}$}	PhaseBal $w = 6$	Stay $w = 1$	$\mathcal{H}$
a. $[_{vP}$ DP $[_{v'}$ v $[_{VP}$ V wh_1]]]	−1		−6
☞ b. $[_{vP}$ wh_1 $[_{v'}$ DP $[_{v'}$ v $[_{VP}$ V t_1]]]]		−1	−1

In a multiple wh-question, the wh-subject is already at the *v*P phase edge and thereby constitutes a potential checker for $C_{[*wh*]}$. Consequently, intermediate movement is not licensed (68b).

(68)

$[_{vP}$ wh_1 $[_{v'}$ v $[_{VP}$ V wh_2]]] ⊕ {T, $C_{[*wh*]}$}	PhaseBal $w = 6$	Stay $w = 1$	$\mathcal{H}$
☞ a. $[_{vP}$ wh_1 $[_{v'}$ v $[_{VP}$ V wh_2]]]			0
b. $[_{vP}$ wh_2 $[_{v'}$ wh_1 $[_{v'}$ v $[_{VP}$ V t_2]]]]		−1	−1

Heck & Müller (2003:109ff.) show that this derives the fact that both multiple wh-fronting and Superiority violations are not possible in languages such as English. Once wh_2 is inside the *v*P, it will not be available at later stages of the derivation due to the PIC. The question now is what parameter will allow us to derive multiple wh-fronting in this system, where a candidate such as (68b) will be selected as optimal. To this end, I adopt the already established view that multiple wh-fronting languages place a wh-probe on the wh-phrases themselves, rather than on C (69b) (see e.g. Bošković 1999, 2002, 2008*a*, 2007; Bailyn 2017).

(69) a. *Single wh-fronting language*:
$[_{\text{CP}}\ \text{C}_{[*\text{wh}*]} \ldots [_{v\text{P}}\ wh_{[\text{wh}]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_{[\text{wh}]}\ ]]]]$
b. *Multiple wh-fronting language*:
$[_{\text{CP}}\ \text{C}_{[\text{wh}]} \ldots [_{v\text{P}}\ wh_{[*\text{wh}*]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ wh_{[*\text{wh}*]}\ ]]]]$

Another important aspect of Heck & Müller's approach to successive-cyclic movement is that final and intermediate steps are driven by different constraints. As Section 5.1 will show, this will allow us to account for the fact that gang effects are triggered relative to Wh-Criterion (at final steps), but not relative to PhaseBalance (at intermediate steps). Whereas simultaneous violations of *Mult-Spec and Lbc were enough to outweigh the violation of Wh-Crit, this is not true of PhaseBalance, which has a higher weight (70).

(70)

$[_{v\text{P}}\ wh_{1[*\text{wh}*]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_{2[*\text{wh}*]}\ \text{NP}]\ ]]]$ $\oplus$ {T, $\text{C}_{[\text{wh}]}$}	PhaseBal $w = 6$	*Mult-Spec $w = 2$	Lbc $w = 2$	$\mathcal{H}$
a. $[_{v\text{P}}\ wh_{1[*\text{wh}*]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_{2[*\text{wh}*]}\ \text{NP}\ ]\ ]]]$	−1			−6
☞ b. $[_{v\text{P}}\ wh_{2[*\text{wh}*]}\ [_{v'}\ wh_{1[*\text{wh}*]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ t_2\ \text{NP}]\ ]]]]$		−1	−1	−4

A further welcome consequence of placing wh-probe features on wh-phrases is that, if a wh-phrase cannot move to its criterial checking position in the specifier of $\text{C}_{[\text{wh}]}$, then its wh-probe feature will remain unchecked. Recall that this was the case for the optimal outputs in the previous analyses of multiple LBE and Superiority. Following (Chomsky 1995, 2000), an unchecked feature is illegible at the interfaces and therefore triggers a crash. This offers then a solution to the notorious problem of ineffability for optimality-theoretic approaches. We can therefore maintain the standard view that the output of the syntactic component may still crash at the interfaces (Müller 2015:897 dubs this the 'bad winners' approach to ineffability).

5 Multiple scrambling and quantifier stranding in Korean

An analogous pattern to the LBE puzzles above can be found with the interaction of multiple scrambling and stranding of numeral quantifiers in Korean. Ko (2007, 2014) shows that there is an incompatibility regarding subject quantifier stranding in multiple scrambling constructions (see Saito 1985; Miyagawa 1989 for similar data from Japanese). First, consider the fact that multiple fronting of a subject QP and an object is generally possible (71) (Ko 2014:45).

(71) [$[_{\text{QP}}$ S Q] O … $[_{v\text{P}}$ $\_\__{\text{QP}}$ $[_{\text{VP}}$ PP $[_{\text{V}'}$ $\_\__{\text{O}}$ V]]]]

$[_{\text{QP}}$ Haksayng-tul-i sey-myeng$]_1$ maykcwu-lul$_2$ $[_{v\text{P}}$ $\_\__1$ kyosil-lo $\_\__2$ kacyewassta]
student-PL-NOM three-CL beer-ACC classroom-to brought
'Three students brought beer to the classroom.'

In addition, Ko (2007, 2014) shows that it is possible for a subject to strand its associated quantifier across a high, propositional adverb such as *pwunmyenghi* ('evidently') (72) (Ko 2014:34).

(72) [S $[_{\text{TP}}$ ADV … $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__{\text{S}}$ Q_{SUB}] O V]]]

Haksayng-tul-i$_1$ $[_{\text{TP}}$ pwunmyenghi $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__1$ sey-myeng] maykcwu-lul masiessta]]
student-PL-NOM evidently three-CL beer-ACC drank
'Evidently, three students drank beer.'

However, if the second step of multiple scrambling involves quantifier stranding from a subject, then it is ruled out (73).

(73) [S O $[_{\text{TP}}$ ADV … $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__{\text{S}}$ Q_{SUB}] PP $\_\__{\text{O}}$ V]]]

?*Haksayng-tul-i$_1$ maykcwu-lul$_2$ $[_{\text{TP}}$ pwunmyenghi $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__1$ sey-myeng] swulcip-eyse $\_\__2$ masiessta]]
student-PL-NOM beer-ACC evidently three-CL bar-LOC drank
'Evidently, three students drank beer.' (Ko 2014:35)

This bears a striking resemblance to the cumulative effect we saw for Slavic, where LBE could not be the second step of multiple wh-fronting. We can therefore view the restriction on multiple scrambling in Korean as an instantiation of the same pattern, but with quantifier stranding instead of LBE. Namely, a single step of scrambling may create a multiple specifier of C or strand a quantifier, but not both simultaneously. This makes the prediction that subject quantifier stranding and multiple fronting should be able to co-occur, if stranding is not the second step. As (74) shows, this prediction is borne out, making this essentially the same kind of Superiority effect we saw with subject LBE in Slavic.

(74) [O S $[_{\text{TP}}$ ADV … $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__{\text{S}}$ Q_{SUB}] PP $\_\__{\text{O}}$ V]]]

Maykcwu-lul$_2$ haksayng-tul-i$_1$ $[_{\text{TP}}$ pwunmyenghi $[_{v\text{P}}$ $[_{\text{QP}}$ $\_\__1$ sey-myeng] swulcip-eyse $\_\__2$ masiessta]]
beer-ACC student-PL-NOM evidently three-CL bar-LOC drank
'Evidently, three students drank beer at a bar.' (Heejeong Ko, p.c.)

At this point, it is important to mention that Ko (2007, 2014) offers a different analysis that also derives these data based on *Cyclic Linearization* (Fox & Pesetsky 2005). I will return to this in Section 5.3. First, let us consider how we can derive the Korean facts in an analogous way to the

Slavic data. Since we are dealing with scrambling, rather wh-movement, the constraint driving movement is the Σ-Criterion in (75). This follows the standard assumption that scrambling is driven by a formal feature [$*\Sigma*$] on the C head (see e.g. Müller 1998; Grewendorf & Sabel 1999; Kawamura 2004; Sabel 2005; Ko 2014).

(75) Σ-Criterion:
XPs bearing [Σ] must be in the specifier of a licensing head bearing [$*\Sigma*$].

Furthermore, stranding can be assumed to violate the constraint *Strand(Q) in (76), against movement that strands a Q head.[6]

(76) *Strand(Q):
Assign a violation for a syntactic object γ in position α in [α … [$_{QP}$ … Q] …], where γ corresponds to β in the input [… [$_{QP}$ β Q] …].

Finally, we also assume the same definition for the constraint *Mult-Spec as in (56). In order to derive the grammatical example in (74), the subject NP must move to Spec-CP first, incurring only a tolerable violation of *Strand(Q) (77b).

(77) *Derivation of* [O S adv … [$_{QP}$ t$_S$ Q] t$_O$] (Step 1):

$[_{CP}\ C_{\substack{[*\Sigma*]\\[*\Sigma*]}}\ \ldots\ [_{vP}\ [_{QP}\ NP_{1[\Sigma]}\ Q\ ]\ \ldots\ NP_2\ ]]$	Σ-Crit $w = 3$	*Mult-Spec $w = 2$	*Strand(Q) $w = 2$	$\mathcal{H}$
a. $[_{CP}\ C_{\substack{[*\Sigma*]\\[*\Sigma*]}}\ \ldots\ [_{vP}\ [_{QP}\ NP_{1[\Sigma]}\ Q\ ]\ \ldots\ NP_{2[\Sigma]}\ ]]$	−2			−6
☞ b. $[_{CP}\ NP_{1[\Sigma]}\ C_{[*\Sigma*]}\ \ldots\ [_{vP}\ [_{QP}\ t_1\ Q\ ]\ \ldots\ NP_{2[\Sigma]}\ ]]$	−1		−1	−5

The following step involves a continuation of (77b). Here, movement of the object NP is possible since it only violates *Mult-Spec (78b).

(78) *Derivation of* [O S adv … [$_{QP}$ t$_S$ Q] t$_O$] (Step 2):

$[_{CP}\ NP_{1[\Sigma]}\ C_{[*\Sigma*]}\ \ldots\ [_{vP}\ [_{QP}\ t_1\ Q\ ]\ \ldots\ NP_{2[\Sigma]}\ ]]$	Σ-Crit $w = 3$	*Mult-Spec $w = 2$	*Strand(Q) $w = 2$	$\mathcal{H}$
a. $[_{CP}\ NP_{1[\Sigma]}\ C_{[*\Sigma*]}\ \ldots\ [_{vP}\ [_{QP}\ t_1\ Q\ ]\ \ldots\ t_2\ ]]$	−1			−3
☞ b. $[_{CP}\ NP_{2[\Sigma]}\ [_{C'}\ NP_{1[\Sigma]}\ C\ \ldots\ [_{vP}\ [_{QP}\ t_1\ Q\ ]\ \ldots\ t_2\ ]]]$		−1		−2

What rules out examples such as (73) is the same as with LBE, violating both *Mult-Spec and *Strand(Q) simultaneously triggers a gang effect. Considering here the crucial final step of multiple fronting in such an example, stranding a quantifier as the second step incurs simultaneous violations of the two constraints and therefore blocks this movement.

[6] This should be viewed as part of a larger 'constraint family' generated by the more general constraint schema *Strand(X). For example, Strand(P) would be the analogous constraint against preposition stranding. This leads to the conclusion that there should be constraints against stranding of other functional heads, as also suggested by Bošković (2016*b*:40).

(79) *Derivation of* $*[\text{ S O ADV} \ldots [_{QP}\ t_S\ Q\ ]\ t_O\ ]$:

$[_{CP}\ NP_{2[\Sigma]}\ C_{[*\Sigma*]} \ldots [_{\nu P}\ [_{QP}\ NP_{1[\Sigma]}\ Q\ ] \ldots t_2\ ]]$	Σ-Crit $w = 3$	*Mult-Spec $w = 2$	*Strand(Q) $w = 2$	$\mathcal{H}$
☞ a. $[_{CP}\ NP_{2[\Sigma]}\ C_{[*\Sigma*]} \ldots [_{\nu P}\ [_{QP}\ NP_{1[\Sigma]}\ Q\ ] \ldots t_2\ ]]$	−1			−3
b. $[_{CP}\ NP_{1[\Sigma]}\ [_{C'}\ NP_{2[\Sigma]}\ C \ldots [_{\nu P}\ [_{QP}\ t_1\ Q\ ] \ldots t_2\ ]]]$		−1	−1	−4

5.1 Subject/object asymmetries

So far, we have focused on cumulative effects involving multiple specifier creation with extraction from a subject (i.e. LBE and Q-stranding). However, comparable effects are absent with object extraction. As Ko (2007:53) shows, an object quantifier can be stranded as the second step of multiple fronting (80).

(80) $[\text{ O S } [_{TP} \text{ ADV} \ldots [_{\nu P}\ \_\__S\ [_{QP}\ \_\__O\ Q_{OBJ}\ ]\ V\ ]]]$

Maykcwu-lul$_2$ haksayng-tul-i$_1$ pwunmyenghi $[_{\nu P}\ [_{QP}\ \_\__1$ sey-pyeng$]\ ___2$ masiessta]
beer-ACC student-PL-NOM evidently three-CL drank
'Evidently, students drank three bottles of beer.'

In fact, this is also a pattern that we find with LBE in Slavic. Recall that it was not possible for subject LBE to be the second step of multiple wh-fronting in Serbo-Croatian (22). However, parallel examples with object LBE as the second step of multiple fronting seem perfectly well-formed (81).[7]

(81) *No Superiority with LBE from object* (Serbo-Croatian):

a. (?)Ko$_1$ kakve$_2$ jasno $\_\__1$ vidi $[_{QP}$ (dve) $[_{NP}\ \_\__2$ devojke]] ?
who what.kind clearly see two girls

b. (?)Kakve$_2$ ko$_1$ jasno $\_\__1$ vidi $[_{QP}$ (dve) $[_{NP}\ \_\__2$ devojke]] ?
what.kind who clearly see two girls
'Who sees what kind of (two) girls clearly?

Thus, objects differ from subjects in that they do not participate in cumulative effects with multiple fronting and sub-extraction. This asymmetry actually supports our current hypothesis that cumulative effects must be triggered by violations local to the same derivational step. A fundamental difference between subjects objects under standard assumptions of phase theory is that the former are base-generated at the edge of the *v*P, whereas objects must first move there to be accessible for subsequent extraction. Under this assumption, sub-extraction from an object incurs the relevant violation at an intermediate step (82a), whereas sub-extraction from a subject

[7] Note that the position of the subject trace in (80) is potentially ambiguous, as being either above or below the adjunct. However, Stjepanović (1999) argues that focused constituents are obligatorily moved to a higher position in the 'middle field' above the subject. If we contrastively focus the adverb *jasno*, this does not affect the overall judgements and we can be confident that the subject has indeed moved.

violates it at the final step (82b).

(82) a. ✓ *Strand(Q)
$[_{\text{CP}}\ \text{NP}_2\ \text{NP}_1 \ldots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ \text{t}_1\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{QP}}\ \text{t}_2\ \text{Q}\ ]\ ]]]]]$
✗ *Strand(Q)

b. $[_{\text{CP}}\ \text{NP}_1\ \text{NP}_2 \ldots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ [_{\text{QP}}\ \text{t}_1\ \text{Q}\ ]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ \text{t}_2\ ]]]]]$
✗ *Strand(Q)

As mentioned in Section 4.4, a gang effect blocks fronting at final, but not intermediate steps, since they are driven by different constraints with different weights. In the derivation of (80), the object NP$_2$ moves to edge of *v*P to provide a potential checker for the second [$*\Sigma*$] on C in the numeration (83b). In doing so, it creates an additional specifier of *v*. However, this is not blocked since the constraint triggering intermediate movement (PhaseBalance) is stronger than the constraint trigger final movement steps (Σ-Crit or Wh-Crit), i.e. it bears a higher weight.[8]

(83) *Derivation of* [O S adv … t$_S$ [$_{\text{QP}}$ t$_O$ Q]] (intermediate step):

$[_{v\text{P}}\ \text{NP}_{1[\Sigma]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{QP}}\ \text{NP}_{2[\Sigma]}\ \text{Q}]\ ]]]$ $\oplus$ {T, $\text{C}_{\left[\substack{*\Sigma* \\ *\Sigma*}\right]}$}	PhaseBal $w = 6$	*Mult-Spec $w = 2$	*Strand(Q) $w = 2$	$\mathcal{H}$
a. $[_{v\text{P}}\ \text{NP}_{1[\Sigma]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{QP}}\ \text{NP}_{2[\Sigma]}\ \text{Q}]\ ]]]$	–1			–6
☞ b. $[_{v\text{P}}\ \text{NP}_{2[\Sigma]}\ [_{v'}\ \text{NP}_{1[\Sigma]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{QP}}\ \text{t}_2\ \text{Q}]\ ]]]]$		–1	–1	–4

The consequence of this is that movement of NP$_2$ from Spec-*v*P no longer violates *Strand(Q). This means it can now create a multiple specifier of C as the second step of multiple fronting since no gang effect will be triggered (84b).

(84) *Derivation of* [O S adv … t$_S$ [$_{\text{QP}}$ t$_O$ Q]] (final step):

$[_{\text{CP}}\ \text{NP}_{1[\Sigma]}\ \text{C}_{[*\Sigma*]} \ldots [_{v\text{P}}\ \text{NP}_{2[\Sigma]}\ [_{v'}\ \text{t}_1 \ldots [_{\text{QP}}\ \text{t}_2\ \text{Q}\ ]\ ]]$	Σ-Crit $w = 3$	*Mult-Spec $w = 2$	*Strand(Q) $w = 2$	$\mathcal{H}$
a. $[_{\text{CP}}\ \text{NP}_{1[\Sigma]}\ \text{C}_{[*\Sigma*]} \ldots [_{v\text{P}}\ \text{NP}_{2[\Sigma]}\ [_{v'}\ \text{t}_1 \ldots [_{\text{QP}}\ \text{t}_2\ \text{Q}\ ]\ ]]$	–1			–3
☞ b. $[_{\text{CP}}\ \text{NP}_{2[\Sigma]}\ [_{\text{C}'}\ \text{NP}_{1[\Sigma]}\ \text{C} \ldots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ \text{t}_1 \ldots [_{\text{QP}}\ \text{t}_2\ \text{Q}\ ]\ ]]]$		–1		–2

The explanation for the lack of Superiority with object LBE is then entirely analogous. The intuition here is that the grammar 'forgets' that the object was sub-extracted since this violation was confined to an earlier step. This a property that I will refer to as *derivational amnesia* and provides a general account of the observed subject/object asymmetries. Finally, intermediate movement to Spec-*v*P with subjects is unmotivated with regard to PhaseBalance and thus, Q-stranding/LBE from a subject takes place directly from its base-position in Spec-*v*P and therefore violates *Strand(Q), as we saw in (79).

[8] The alternative approach would be to say that intermediate steps always precede merger of the subject and thus form the first specifier (cf. *Intermediate Step Corollary*; Müller 2011:176). This can be determined by the relative weight of constraints of PhaseBalance and the constraint triggering external (e.g. MergeCondition; Heck & Müller 2013:138), see Section 5.2.

5.2 A potential loophole

While the assumption that sub-extracted objects can avoid gang effects by incurring extraction-related violations at an intermediate step provides an account of subject/object asymmetries, it potentially undermines the previous account of the ban on multiple LBE. The reason for this is that it introduces a loophole whereby object LBE can avoid a cumulative effect at Spec-CP by violating LBC at an intermediate step. In (85), movement of wh_2 should be possible as the second step of multiple fronting, since it only violates *MULT-SPEC but not LBC.

(85) ✗LBC ✗*MULT-SPEC

$[_{\text{CP}}\ \underbrace{wh_2\ [_{\text{C}'}\ wh_1\ \text{C}_{[\text{wh}]} \ldots\ [_{v\text{P}}\ \text{t}_2}_{✓\text{LBC}\ ✗\text{*MULT-SPEC}}\ [_{v'}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

The important thing is the unwanted derivation in (85) requires that the subject left-branch wh_1 moves as the first step of multiple fronting in order to avoid a gang effect by creating a multiple specifier. Thus, the key to ruling out (85) lies in restricting movement of the subject left-branch if the intermediate step involved LBE. The first part of the explanation involves the assumption that, ordinarily, specifiers of v can be created in either order. In other words, it is possible to either merge the subject or perform intermediate movement of the object as the first specifier of v. This can be viewed as the result of the constraint driving intermediate movement (PHASEBALANCE) and the constraint triggering external merge of the subject (MERGECONDITION) bearing the same weight and therefore both being equally available options. However, (86) shows that, if the intermediate step also involves LBE, this tie is broken by the additional violation of LBC (86c) and merger of the subject will be preferred (86b).

(86)

$[_{v\text{P}}\ v_{[\bullet\text{N}\bullet]}\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_{2[*\text{wh}*]}\ \text{NP}]\ ]]$ $\oplus \{wh_{1[*\text{wh}*]}, \text{T}, \text{C}_{[\text{wh}]}\}$	PHASEBAL $w = 6$	MERGECON $w = 6$	LBC $w = 2$	$\mathcal{H}$
a. $[_{v\text{P}}\ v_{[\bullet\text{N}\bullet]}\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_{2[*\text{wh}*]}\ \text{NP}]\ ]]$	−1	−1		−12
☞ b. $[_{v\text{P}}\ wh_{1[*\text{wh}*]}\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ wh_{2[*\text{wh}*]}\ \text{NP}]\ ]]]$	−1			−6
c. $[_{v\text{P}}\ wh_{2[*\text{wh}*]}\ [_{v'}\ v_{[\bullet\text{N}\bullet]}\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]$		−1	−1	−8

The consequence of this is that the subject must necessarily be merged first as the inner specifier if the intermediate step involves LBE. This now interacts with another restriction on movement. Generally, let us assume that multiple specifiers of v count as equidistant to a higher landing site, such that extraction of either the subject or object can take place as the first step (accounting for the basic the lack of Superiority). However, I propose the following constraint, stating that sub-extraction must take place from the outermost available specifier:[9]

[9] In Murphy (2017), I show how this condition can be derived as a cumulative effect between a general (violable) constraint on extraction from inner edges (Bošković 2016*a*) and LBE. For present purposes, I will simply adopt it as a theoretical postulate in itself.

(87) *Condition on extraction from recursive edges*:
Extraction from a recursive edge (i.e. a specifier of α in a specifier of β) is only possible if β is the outermost specifier.
$*[\text{ YP} \dots [_{\beta\text{P}}\ \text{XP}\ [_{\beta'}\ [_{\alpha\text{P}}\ \text{t}_{\text{YP}}\ [_{\alpha'} \dots]]\ [_{\beta'}\ \beta \dots]]$

In other words, LBE must take place from the outer specifier of the v. These two independent conditions work together to close the aforementioned loophole in the following way: intermediate movement cannot target an inner specifier if it is LBE (86). Thus, the a complex subject will always have to occupy the inner specifier of v, a position from which LBE is not possible given (87). Considering the four logically possible derivations for multiple LBE, (88c,d) are both ruled because intermediate LBE targets the inner specifier of v. Furthermore, both (88a) and (88b) are ruled out by (87), since they involve extraction from an inner specifier (in fact (88a) also involves LBE as the second step of multiple fronting).

(88) *Multiple LBE not possible*:

a. $*[_{\text{CP}}\ wh_1\ wh_2\ \text{C}_{\text{wh}} \dots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

b. $*[_{\text{CP}}\ wh_2\ wh_1\ \text{C}_{\text{wh}} \dots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

c. $*[_{\text{CP}}\ wh_1\ wh_2\ \text{C}_{\text{wh}} \dots [_{v\text{P}}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ \text{t}_2\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

d. $*[_{\text{CP}}\ wh_2\ wh_1\ \text{C}_{\text{wh}} \dots [_{v\text{P}}\ [_{\text{NP}}\ \text{t}_1\ \text{NP}]\ [_{v'}\ \text{t}_2\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

Consequently, there is no way of deriving the unattested pattern of multiple LBE. Nevertheless, we still derive the correct outcome for mixed multiple fronting with object LBE. Recall that in such cases, there was no Superiority restriction and both orders of extraction were equally possible. Since intermediate LBE must target the outer specifier, the derivations in (89c) and (89d) are ruled out. However, since subject extraction does not involve LBE, the condition in (87) is not relevant. Consequently, extraction in either order is possible (89a,b), as with ordinary wh-questions.

(89) *Object LBE possible with both orders*:

a. $[_{\text{CP}}\ wh_1\ wh_2\ \text{C}_{\text{wh}} \ldots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ \text{t}_1\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

b. $[_{\text{CP}}\ wh_2\ wh_1\ \text{C}_{\text{wh}} \ldots [_{v\text{P}}\ \text{t}_2\ [_{v'}\ \text{t}_1\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

c. $^*[_{\text{CP}}\ wh_1\ wh_2\ \text{C}_{\text{wh}} \ldots [_{v\text{P}}\ \text{t}_1\ [_{v'}\ \text{t}_2\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

d. $^*[_{\text{CP}}\ wh_2\ wh_1\ \text{C}_{\text{wh}} \ldots [_{v\text{P}}\ \text{t}_1\ [_{v'}\ \text{t}_2\ [_{v'}\ v\ [_{\text{VP}}\ \text{V}\ [_{\text{NP}}\ \text{t}_2\ \text{NP}]\ ]]]]]$

As for subject LBE, extraction will only be possible from the outermost specifier of *v* and will have to respect the condition that LBE cannot be the second step of multiple fronting (leading to the Superiority effect we saw previously). Thus, the assumption that both intermediate steps of LBE and sub-extraction from inner specifiers are restricted work together to rule out multiple LBE, while still accounting for the relevant subject/object asymmetry.

5.3 Against a Cyclic Linearization alternative

Let us now briefly compare this to Ko's (2007; 2014) linearization-based account. Fox & Pesetsky's (2005) theory of *Cyclic Linearization* is based on the idea that successive-cyclic movement is driven by the need to avoid contradictory linearization statements between Spell-Out domains. Ko (2007, 2014) argues that the ban on subject stranding (73) (repeated schematically below) can be accounted for in this way.

(90) ?*[S O $[_{\text{TP}}$ ADV … $[_{v\text{P}}\ [_{\text{QP}}\ \text{t}_{\text{S}}\ \text{Q}_{\text{SUB}}\ ]\ \text{t}_{\text{O}}\ \text{V}\ ]]]$

For the ungrammatical (90a), there are two possible derivations. The first involves no movement of the object to Spec-*v*P. This generates the linearization statement that the subject quantifier precedes the object $Q_{SUB} < O$ (91a). At the CP-level, extraction of the object followed by the subject gives rise to the conflicting linearization instruction that the object precedes the subject quantifier ($O < Q_{SUB}$) (91b), leading to ungrammaticality.

(91) *Option 1 (no object scrambling)*:

a. $[_{v\text{P}}\ [_{\text{QP}}\ \text{S}\ \text{Q}_{\text{SUB}}\ ]\ [_{\text{VP}}\ \text{O}\ \text{V}\ ]\ v\ ]$
Ordering at vP: S < $\mathbf{Q_{SUB}}$ < **O** < V < *v*

b. $^*[_{\text{CP}}\ \text{S}\ \text{O}\ [_{\text{TP}}\ \text{ADV}\ [_{v\text{P}}\ [_{\text{QP}}\ \text{t}_{\text{S}}\ \text{Q}_{\text{SUB}}\ ]\ [_{\text{VP}}\ \text{t}_{\text{O}}\ \text{V}\ ]\ v\ ]\ \text{T}\ ]\ \text{C}\ ]$
Ordering at CP: S < **O** < ADV < $\mathbf{Q_{SUB}}$ < V < *v* < T < C

The problematic linearization statement $Q_{SUB} < O$ can be avoided be moving the object in front of the subject quantifier at the *v*P level (92a). However, this now creates a different problem. The object now precedes the subject inside the *v*P Spell-Out domain, meaning that this order cannot be reversed at Spec-CP (92b).

(92) *Option 2 (with object scrambling)*:

a. $[_{vP}$ O $[_{QP}$ S Q_{SUB}] $[_{VP}$ t_O V] v]
Ordering at vP: **O** < **S** < Q_{SUB} < V < v

b. *$[_{CP}$ S O $[_{TP}$ ADV $[_{vP}$ t_O $[_{QP}$ t_S Q_{SUB}] $[_{VP}$ t_O V] v] T] C]
Ordering at CP: **S** < **O** < ADV < Q_{SUB} < V < v < T < C

The derivation in (92) should be possible if the subject and object are fronted in the reverse order. Recall that this was shown to be grammatical in (74). Thus, Ko's (2007; 2014) analysis can derive the ban on subject stranding as the second step of multiple fronting with reference to linearization. Indeed, we can derive the Superiority restriction with subject LBE in Slavic, summarized in (93), in a similar way.

(93) a. *$[_{CP}$ wh_1 wh_2 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V t_2]]]]]

b. $[_{CP}$ wh_2 wh_1 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V t_2]]]]]

The object wh_2 first has to move to the edge of *v*P to avoid a later contradiction relative to NP (parallel to Q_{sub} above). this creates the linearization statement $wh_2 < wh_1$ (94a). Consequently, the order of the wh-phrases at Spec-CP must respect this ordering. This rules out (94b), but allows for the derivation in (94b′).

(94) *Cyclic Linearization analysis of subject LBE*:

a. $[_{vP}$ wh_2 $[_{v'}$ $[_{NP}$ wh_1 NP] $[_{v'}$ v $[_{VP}$ V t_2]]]]
Ordering at vP: $\mathbf{wh_2}$ < $\boldsymbol{wh_1}$ < NP < v < V

b. *$[_{CP}$ wh_1 wh_2 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V t_2]]]]]
Ordering at CP: $\mathbf{wh_1}$ < $\boldsymbol{wh_2}$ < NP < v < V

b.′ $[_{CP}$ wh_2 wh_1 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V t_2]]]]]
Ordering at CP: $\mathbf{wh_2}$ < $\boldsymbol{wh_1}$ < NP < v < V

While Ko's Cyclic Linearization approach can equally account for Superiority with subject LBE, recall that this was only one half of the data that the alternative cumulative analysis could account for. In addition, we saw that ruling out LBE as the second step of multiple fronting also provides an analysis of the ban on multiple LBE (95).

(95) a. *$[_{CP}$ wh_1 wh_2 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V $[_{NP}$ t_2 NP]]]]]]

b. *$[_{CP}$ wh_2 wh_1 … $[_{vP}$ t_2 $[_{v'}$ $[_{NP}$ t_1 NP] $[_{v'}$ v $[_{VP}$ V $[_{NP}$ t_2 NP]]]]]]

This is not predicted by the Cyclic Linearization analysis. In (96), the only difference to the analysis in (94) is the presence of the additional linearization statement $V < NP$ at the *v*P level. The CP order in (96b′) does not contradict any linearization statements from *v*P and is therefore predicted to be grammatical, contrary to fact.

(96) *Cyclic Linearization analysis of multiple LBE*:

a. [$_{vP}$ wh_2 [$_{v'}$ [$_{NP}$ wh_1 NP] [$_{v'}$ v [$_{VP}$ V [$_{NP}$ t_2 NP]]]]]
Ordering at vP: $\mathbf{wh_2} \prec \boldsymbol{wh_1} \prec \text{NP} \prec v \prec \text{V} \prec \text{NP}$

b. *[$_{CP}$ wh_1 wh_2 … [$_{vP}$ t_2 [$_{v'}$ [$_{NP}$ t_1 NP] [$_{v'}$ v [$_{VP}$ V [$_{NP}$ t_2 NP]]]]]]
Ordering at CP: $\mathbf{wh_1} \prec \boldsymbol{wh_2} \prec \text{NP} \prec v \prec \text{V} \prec \text{NP}$

b.′ [$_{CP}$ wh_2 wh_1 … [$_{vP}$ t_2 [$_{v'}$ [$_{NP}$ t_1 NP] [$_{v'}$ v [$_{VP}$ V [$_{NP}$ t_2 NP]]]]]]
Ordering at CP: $\mathbf{wh_2} \prec \boldsymbol{wh_1} \prec \text{NP} \prec v \prec \text{V} \prec \text{NP}$

Consequently, it seems that the two analyses are not simply notational variants, and that the present analysis can achieve better explanatory coverage of the more complex LBE facts than its Cyclic Linearization alternative.

6 Multiple correlative displacement in Hindi

The final cumulative effect with multiple fronting that we will discuss involves correlative constructions in Hindi (also see Srivastav 1991; Dayal 1996; Mahajan 2000; Bhatt 2003, 2015). Several languages have a correlativization strategy for nominal modification (see Lipták 2009 for an overview). The basic structure of a correlative in Hindi involves a left-peripheral (free) relative clause, which is co-indexed with a demonstrative phrase in the matrix clause (97).

(97) [jo CD sale-par hai]$_i$ Maya [us CD-ko]$_i$ khari:d-egi:
REL CD sale-on be.PRES Maya DEM CD-ACC buy-FUT.F
'Maya will buy the CD that is on sale.'
(Lit. [Which CD is one sale], Maya will buy that CD) (Bhatt 2003:486)

Although this looks like a non-local dependency, Bhatt (2003) has demonstrated that the relation between the correlative clause and the demonstrative is actually one of movement. First, consider the fact the dependency between the correlative and the demonstrative can span a finite clause boundary (98) (note that Bhatt 2003:500 also shows that binding is not sensitive to islands).

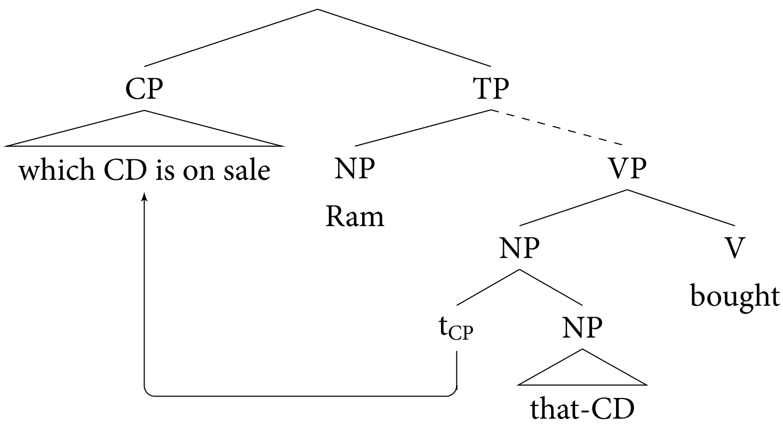
(98) [$_{CP}$ jo larki: TV-par ga: rah-i: hai]$_i$ [$_{CP}$ Sita soch-ti: hai [$_{CP}$ ki vo$_i$
REL girl TV-on sing PROG be.PRES Sita think-HAB.F be.PRES that DEM
sundar hai]]
beautiful BE.PRES
'Sita thinks that the girl who is singing on TV is beautiful.' (Bhatt 2003:500)

However, the correlative CP cannot be separated from the demonstrative by an island boundary, for example, as shown with the Complex NP Island in (99).

(99) *[$_{CP}$ jo vahã rah-ta: hai]$_i$ muhj-ko [$_{NP}$ vo kaha:ni: [$_{CP}$ jo Arundhati-ne
REL there stay-HAB be.PRES 1SG-DAT that story.F REL Arundhati-ERG
us-ke-baare-mẽ$_i$ likh-ii]] pasand hai
DEM-about write.PERF.F like be.PRES
'Who lives there$_i$, I like the story that Arundhati wrote about that boy.'
(Bhatt 2003:500)

The fact that the dependency between the CorCP and the Dem-XP is unbounded and constrained by islands leads Bhatt (2003) to the conclusion that the CP is base-generated as an adjunct to the demonstrative phrase and subsequently displaced to a higher position (100).

(100) *Structure of Hindi correlatives* (Bhatt 2003:497):



With this structure in mind, consider the fact that it is also possible to have multiple correlative clauses in a single clause, adjacent to their demonstrative associates (101).

(101) $[_{CP}$ … [CorCP$_1$ Dem-XP$_1$] … [CorCP$_2$ Dem-XP$_2$] …]
Ram-ne $[_{NP}$ $[_{CP}$ jo laṛkaa tumhaare pi:chhe hai $]_1$ $[_{NP}$ us laṛke-ko $]_1$] $[_{NP}$ $[_{CP}$
Ram-ERG REL boy your behind be.PRES DEM boy-DAT
jo kita:b Shantiniketan-ne chhaapii thii $]_2$ $[_{NP}$ vo kitaab $]_2$] dii
REL book Shantiniketan-ERG print.PERF.F was.F DEM book give.PERF-F
'Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. 'Ram gave [[which book Shantiniketan had published] that book] to [[which boy is behind you] that boy]')

(Bhatt 2003:507)

Furthermore, Bhatt (2003) shows that is possible to front one of these correlative clauses, either CorCP$_1$ associated with the indirect object (102), or CorCP$_2$ modifying the direct object (103).

(102) $[_{CP}$ CorCP$_1$ … [t_{CorCP_1} Dem-XP$_1$] … [CorCP$_2$ Dem-XP$_2$] …]
$[_{CP}$ jo laṛkaa tumhaare pi:chhe hai $]_1$ Ram-ne $[_{NP}$ t_{CP} $[_{NP}$ us laṛke-ko]]$_1$ $[_{NP}$
REL boy your behind be.PRES Ram-ERG DEM boy-DAT
$[_{CP}$ jo kita:b Shantiniketan-ne chhaapii thii $]_2$ $[_{NP}$ vo kitaab $]_2$] dii
REL book Shantiniketan-ERG print.PERF.F was.F DEM book give.PERF-F
'Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which boy is behind you] Ram gave [[which book Shantiniketan had published] that book] to [that boy]')

(Bhatt 2003:507)

(103) $[_{CP}$ CorCP$_2$ … [CorCP$_1$ Dem-XP$_1$] … [t_{CorCP_2} Dem-XP$_2$] …]
$[_{CP}$ jo kita:b Shantiniketan-ne chhaapii thii $]_2$ Ram-ne $[_{NP}$ $[_{CP}$ jo laṛkaa
REL book Shantiniketan-ERG print.PERF.F was.F Ram-ERG REL boy

tumhaare pi:chhe hai]$_1$ [$_{NP}$ us laṛke-ko]$_1$] [$_{NP}$ t$_{CP}$ [$_{NP}$ vo kitaab]$_2$] dii
your behind be.PRES DEM boy-DAT DEM book give.PERF-F
'Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which book Shantiniketan had published] Ram gave [that book] to [[which boy is behind you] that boy]')

(Bhatt 2003:507)

In addition, Hindi generally allows for long-distance scrambling of multiple constituents to a left-peripheral position. This is shown for DPs arguments and wh-phrases in (104).

(104) a. [$_{CP}$ Ram-ne$_1$ Sita-ko$_2$ Radha soch-ti: hai [$_{CP}$ ki t$_1$ t$_2$ kai tohfe
Ram-ERG Sita-DAT Radha think-HAB.F be that many presents
di-ye the]]
give-PERF.PL be.PAST.MPL
'Radha thinks that Ram gave Sita many presents.'

b. [$_{CP}$ kis-ne$_1$ kis-ko$_2$ Radha soch-ti: hai [$_{CP}$ ki t$_1$ t$_2$ kai tohfe
who-ERG who-DAT Radha think-HAB.F be that many presents
di-ye the]]
give-PERF.PL be.PAST.MPL
'Radha thinks that who gave whom many presents?' (Bhatt 2003:509)

Interestingly, it is not possible to have multiple fronting of correlative clauses, where both CorCP$_1$ and CorCP$_2$ are moved to clause-initial position (105) (Bhatt 2003:508). As (106) shows, the order of the fronted CPs does not make a difference.

(105) *[$_{CP}$ CorCP$_1$ CorCP$_2$ … [t$_{CorCP_1}$ Dem-XP$_1$] … [t$_{CorCP_2}$ Dem-XP$_2$] …]

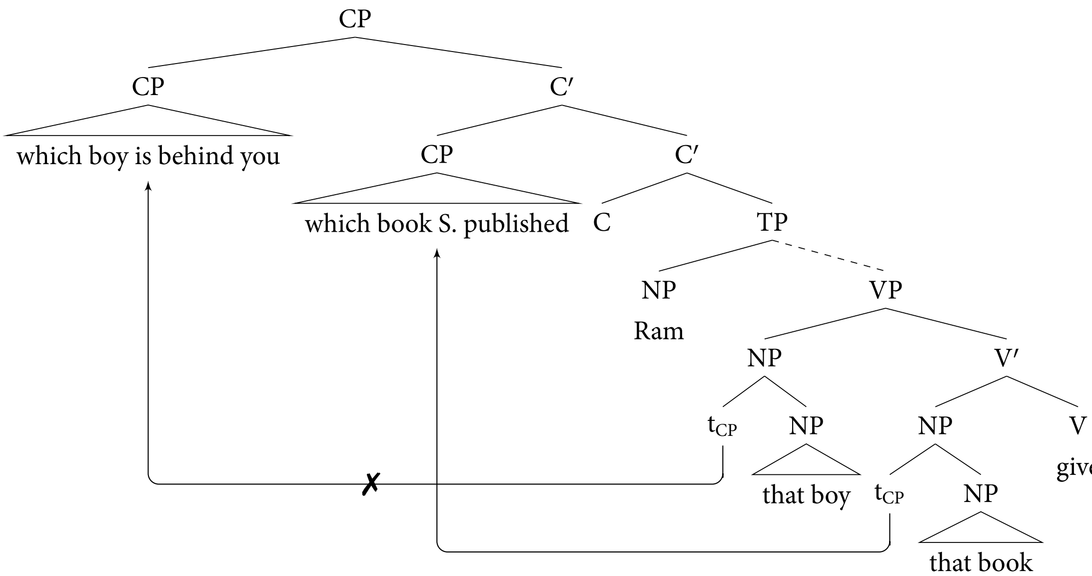
*[$_{CP}$ jo laṛkaa tumhaare pi:chhe hai]$_1$ [$_{CP}$ jo kita:b Shantiniketan-ne chhaapii
REL boy your behind be REL book Shantiniketan-ERG print.PERF.F
thii]$_2$ Ram-ne [$_{NP}$ t$_{CP}$ [$_{NP}$ us laṛke-ko]$_1$] [$_{NP}$ t$_{CP}$ [$_{NP}$ vo kitaab]$_2$] dii
was.F Ram-ERG DEM boy-DAT DEM book give.PERF-F
'Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which boy is behind you] [which book Shantiniketan had published] Ram gave [that book] to [that boy]')

(106) *[$_{CP}$ CorCP$_2$ CorCP$_1$ … [t$_{CorCP_1}$ Dem-XP$_1$] … [t$_{CorCP_2}$ Dem-XP$_2$] …]

*[$_{CP}$ jo kita:b Shantiniketan-ne chhaapii thii]$_2$ [$_{CP}$ jo laṛkaa tumhaare pi:chhe
REL book Shantiniketan-ERG print.PERF.F was.F REL boy your behind
hai]$_1$ Ram-ne [$_{NP}$ t$_{CP}$ [$_{NP}$ us laṛke-ko]$_1$] [$_{NP}$ t$_{CP}$ [$_{NP}$ vo kitaab]$_2$] dii
be.PRES Ram-ERG DEM boy-DAT DEM book give.PERF-F
'Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which book Shantiniketan had published] [which boy is behind you] Ram gave [that book] to [that boy]')

We can conceive of the impossibility of multiple correlative displacement as a cumulative effect in the following way: In general, Hindi allows for a correlative CP to be fronted. Furthermore, it also possible to have multiple fronting of XPs (104). However, the combination of these two processes, i.e. multiple fronting of correlative CPs, is not permitted (107).

(107) *No multiple fronting of Hindi correlatives*:



In the same vein as preceding analysis for Serbo-Croatian and Korean, let us assume that multiple fronting violates the constraint *Mult-Spec. The question is now what constraint militates again correlative fronting. As (107) shows, Bhatt (2003) argues that correlative fronting is actually movement of an adjunct. We can therefore assume that there is a violable constraint against movement of adjoined phrases, *Move(Adjunct) (108), which is violated by correlative fronting.

(108) *MoveAdj(unct):
Movement of an adjunct is prohibited.

This is supported by evidence in Bhatt (2003:509f.) showing that, while verbal adjuncts can generally undergo long-distance movement (109a), they cannot undergo multiple fronting (109b) (unlike arguments, cf. (104)).

(109) a. [$_{CP}$ kab$_1$ Radha soch-ti: hai [$_{CP}$ ki Ram-ne Sita-ko tohfe t_1
when Radha think-HAB.F be that Ram-ERG Sita-DAT presents
di-ye the]] ?
give-PERF.PL be.PAST.MPL
'When does Radha think that Ram gave presents to Sita?'

b. *[$_{CP}$ kab$_2$ kahã$_1$ Radha soch-ti: hai [$_{CP}$ ki Ram-ne Sita-ko tohfe t_1 t_2
when where Radha think-HAB.F be that Ram-ERG Sita-DAT presents
di-ye the]] ?
give-PERF.PL be.PAST.MPL
'When and where does Radha think that Ram gave presents to Sita?'

(Bhatt 2003:510)

The fact that adjuncts behave in the same way as correlatives in not being able to undergo multiple fronting supports the assumption that (108) is relevant constraint.[10] We can therefore analyze this in an entirely analogous fashion to the Korean and Slavic data. The respective weights of the constraints against multiple fronting (*Mult-Spec) and correlative fronting (*Move(Adjunct)) are individually lower than constraint driving scrambling (Σ-Criterion), however their summed weights are higher. Considering the relevant step of the derivation in which the first correlative clause is fronted, this step is harmonically-improving and therefore licensed (110b). [11]

(110) *Step Σ*:

$[_{CP}\ C_{\left[\substack{*\Sigma* \\ *\Sigma*}\right]} \ldots [\ CorCP_{1[\Sigma]}\ NP_1\ ] \ldots [\ CorCP_{2[\Sigma]}\ NP_2\ ]\ ]$	Σ-Crit $w = 3$	*Move(Adj) $w = 2$	*Mult-S $w = 2$	$\mathcal{H}$
a. $[_{CP}\ C_{\left[\substack{*\Sigma* \\ *\Sigma*}\right]} \ldots [\ CorCP_{1[\Sigma]}\ NP_1\ ] \ldots [CorCP_{2[\Sigma]}\ NP_2\ ]\ ]$	−2			−6
☞ b. $[_{CP}\ CorCP_1\ C_{[*\Sigma*]} \ldots [\ t_1\ NP_1\ ]] \ldots [\ CorCP_{2[\Sigma]}\ NP_2\ ]\ ]$	−1	−1		−5

At the subsequent step, we try to move the second correlative clause, however this step simultaneously involves adjunct movement and creation of multiple specifier, which thereby triggers a gang effect (111b). As a result, this movement step is blocked.

(111) *Step Σ_{+1}*:

$[_{CP}\ CorCP_2\ C_{[*\Sigma*]} \ldots [\ CorCP_{1[\Sigma]}\ NP_1\ ] \ldots [\ t_2\ NP_2\ ]\ ]$	Σ-Crit $w = 3$	*Move(Adj) $w = 2$	*Mult-S $w = 2$	$\mathcal{H}$
☞ a. $[_{CP}\ CorCP_2\ C_{[*\Sigma*]} \ldots [\ CorCP_{1[\Sigma]}\ NP_1\ ] \ldots [\ t_2\ NP_2\ ]\ ]$	−1			−3
b. $[_{CP}\ CorCP_{1[\Sigma]}\ [_{C'}\ CorCP_2\ C \ldots [\ t_1\ NP_1\ ] \ldots [\ t_2\ NP_2\ ]\ ]]$		−1	−1	−4

This analysis is further supported by the novel observation that multiple fronting of correlative clauses is possible if their criterial positions are in different clauses (Rajesh Bhatt, p.c.). If one of the CorCPs moves is scrambled to a higher clause than the other, then multiple displacement is possible, as (112) and (113) show.

[10] Further evidence for this comes from the fact that multiple fronting of adjuncts is also reported to be ungrammatical in Slavic languages (i). This suggests that a similar restriction on multiple specifier creation by adjunct movement can also hold at final steps.

(i) a. *[$_{CP}$ Dlaczego$_1$ kiedy$_2$ [$_{TP}$ wyjechałeś __$_1$ __$_2$ z kraju]] ?
why when leave.2sg out.of country
'When and why did you leave the country?' (*Polish*; Cichocki 1983:56)

b. *[$_{CP}$ Gdje$_1$ kada$_2$ [$_{TP}$ Ivan nastupa __$_1$ __$_2$]] ?
where when Ivan performs
'Where and when does Ivan perform?'
(*Serbo-Croatian*; Citko & Gračanin-Yuksek 2012:24)

[11] I am not considering the *v*P phase here. It may well be the case that the correlative 'big NP' first moves to the edge of *v*P before sub-extraction of CorCP takes place. However due to Hindi being a head-final language, it is notoriously difficult to diagnose the height of elements inside the *v*P. Furthermore, Keine (2016:140ff.) presents an analysis of scrambling in Hindi that he shows to be incompatible with the existence of *v*P phases in the language. In light of this, I will remain agnostic about the issue.

(112) $[_{CP_2}$ CorCP$_1$ I think that $[_{CP_1}$ CorCP$_2$ … [t_{CorCP_1} Dem-XP$_1$] … [t_{CorCP_2} Dem-XP$_2$] …]]

$[_{CP}$ jo laṛkaa tumhaare pi:chhe hai $]_1$ mujhe lagtaa hai ki $[_{CP}$ jo
REL boy your behind be.PRES be.PRES me.DAT feel.HAB be.PRES REL
kita:b Shantiniketan-ne chhaapii thii $]_2$ Ram-ne $[_{NP}$ t_{CP} $[_{NP}$ us laṛke-ko $]_1$]
book Shantiniketan-ERG print.PERF.F was.F Ram-ERG DEM boy-DAT
$[_{NP}$ t_{CP} $[_{NP}$ vo kitaab $]_2$] dii
DEM book give.PERF-F

'I think that Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which boy is behind you] I think that [which book Shantiniketan had published] Ram gave [that book] to [that boy]')

(Rajesh Bhatt, p.c.)

(113) $[_{CP}$ CorCP$_2$ I think that $[_{CP}$ CorCP$_1$ … [t_{CorCP_1} Dem-XP$_1$] … [t_{CorCP_2} Dem-XP$_2$] …]]

$[_{CP}$ jo kita:b Shantiniketan-ne chhaapii thii $]_2$ mujhe lagtaa hai ki
REL book Shantiniketan-ERG print.PERF.F was.F be.PRES me.DAT feel.HAB be.PRES
$[_{CP}$ jo laṛkaa tumhaare pi:chhe hai $]_1$ Ram-ne $[_{NP}$ t_{CP} $[_{NP}$ us laṛke-ko $]_1$] $[_{NP}$
REL boy your behind be.PRES Ram-ERG DEM boy-DAT
t_{CP} $[_{NP}$ vo kitaab $]_2$] dii
DEM book give.PERF-F

'I think that Ram gave the book that Shantiniketan had published to the boy who is standing behind you.' (Lit. '[which boy is behind you] I think that [which book Shantiniketan had published] Ram gave [that book] to [that boy]')

(Rajesh Bhatt, p.c.)

This follows from the previously established assumption that gang effects involving multiple specifier creation can hold at final steps, but not intermediate steps. [12] Recall that this was due to the fact that intermediate steps are driven by a different constraint with a possibly different weight, namely PHASEBALANCE. This means that a multiple specifier of C$_1$ is licensed by an intermediate step (but not a final step), as in (114).

(114) $[_{CP_2}$ CorCP$_1$ I think that $[_{CP_1}$ t_1 $[_{C'_1}$ CorCP$_2$ … [t_{CorCP_1} NP$_1$] … [t_{CorCP_2} NP$_2$]]]]

As in previous cases, this is because PHASEBALANCE bears a higher weight than Σ-CRITERION and is therefore immune from the cumulative effect of summed violations. Consider the step of

[12] It is conceivable that some languages have weighting conditions that would lead to a gang effect at an intermediate step. One plausible example of this involves the 'selective' nature of wh-islands (Postal 1998). The well-known argument/adjunct asymmetry in (i) could be explained under the assumption that multiple specifier creation by adjunct movement is ruled out at *intermediate* steps in English. The wh-phrase extracted from a wh-island passes through a second specifier of C. This is possible for a wh-argument (ia), but not for a wh-adjunct (ib).

(i) a. What$_1$ do you wonder $[_{CP}$ t_1 $[_{C'}$ how$_2$ to repair t_1 t_2]] ?

b. *How$_2$ do you wonder $[_{CP}$ t_2 $[_{C'}$ what$_1$ to repair t_1 t_2]] ? (Manzini 1997:135f.)

the derivation in which one correlative clause has moved to Spec-CP. The movement step in (115b) is licensed to provide a potential checker for the [∗Σ∗] feature on the C head in the numeration.

(115)

$[_{CP}$ CorCP$_2$ C … [CorCP$_{1[\Sigma]}$ NP$_1$] … [t$_2$ NP$_2$]] ⊕ {I, think, v, T, C$_{[*\Sigma*]}$}	PhaseBal $w = 6$	*Mv(Adj) $w = 2$	*Mult-S $w = 2$	$\mathcal{H}$
a. $[_{CP}$ CorCP$_2$ C … [CorCP$_{1[\Sigma]}$ NP$_1$] … [t$_2$ NP$_2$]]	−1			−6
☞ b. $[_{CP}$ CorCP$_{1[\Sigma]}$ $[_{C'}$ CorCP$_2$ C … [t$_1$ NP$_1$] … [t$_2$ NP$_2$]]]		−1	−1	−4

Thus, the case of Hindi correlative fronting provides another example of restricted multiple specifier creation. Furthermore, the initially surprising fact that multiple fronting is possible when the CorCPs do not land in the same clause provides another example that cumulative effects at final and intermediate movement steps, as predicted by an analysis in which the two steps are driven by different constraints.

7 Conclusion

This paper has argued that certain restrictions found with multiple fronting are the result of cumulative constraint interaction. Three cases studies were provided involving Left-Branch Extraction in Slavic, quantifier stranding in Korean and correlative displacement in Hindi. It was claimed that there is a general, violable constraint in the grammar against representations containing multiple specifiers of a single head (*Mult-Spec). In each of the aforementioned case studies, violations of *Mult-Spec are tolerable in isolation, leading to the possibility of multiple fronting in these languages, however not in conjunction with another violation incurred by a marked extraction process. The abstract pattern underlying these three cases can be abstractly summarized in (116).

(116) $^*[_{ZP}$ XP $[_{Z'}$ YP $[_{Z'}$ Z … [… $[_{NP}$ … __$_{XP}$ …] … __$_{YP}$ …]]]]

It was shown that a language can avoid the banned configuration in (116) by having the movement step involving sub-extraction apply as the first step, thereby spreading the respective violations out across different steps of multiple fronting. This gave rise to exceptional ordering restrictions with both LBE in Slavic and quantifier stranding in Korean of the kind not otherwise found in the languages. However, an important caveat is that these were only shown to hold for subject extraction. It was argued that this reveals a fundamental asymmetry between subject and objects, namely that extraction from an object first involves intermediate movement. It is this property that allows for the circumvention of gang effects by again distributing the violations across intermediate and final movement steps. Furthermore, the fact that cumulative blocking is triggered at final, but not intermediate steps lends support to the existing claim by Heck & Müller (2003) that these are driven by distinct constraints, which can in turn bear different weights. It was also shown that cumulative constraint violations must occur local to the same derivational step, which provides a strong argument for a local, derivational approach such as Serial Harmonic Grammar. While this may seem like an enrichment to existing Minimalist theories, it is arguably inevitably required in any sufficiently explicit theory of local economy and cumulativity. Given the undeniable similarities between the three case studies, an approach such as Serial Harmonic Grammar provides a theory that can directly capture the core intuition that we are dealing with the illicit combination of ordinarily licit process in the languages in question.

References

Ackema, Peter & Ad Neeleman (1998). Optimal Questions. *Natural Language and Linguistic Theory* 16(3). 443–490.

Adli, Aria (2011). A Heuristic Mathematical Approach for Modeling Constraint Cumulativity: Contrastive Focus in Spanish and Catalan. *The Linguistic Review* 28(2). 111–173.

Avgustinova, Tania (1994). On Bulgarian Verb Clitics. *Journal of Slavic Linguistics* 2(1). 29–47.

Bailyn, John F. (2017). Bulgarian Superiority and Minimalist Movement Theory. In Y. Oseki, M. Esipova & S. Harves (eds). *Proceedings of Formal Approaches to Slavic Linguistics 24: The NYU Meeting*. Michigan Slavic Publications: Ann Arbor, MI. 27–49.

Bhatt, Rajesh (2003). Locality in Correlatives. *Natural Language and Linguistic Theory* 21(3). 485–541.

Bhatt, Rajesh (2015). Relative Clauses and Correlatives. In A. Alexiadou & T. Kiss (eds). *Syntax – Theory and Analysis: An International Handbook*. de Gruyter: Berlin. 708–749.

Billings, Loren (2002). Why Clitics Cluster Together in Balkan Slavic: Non-Templatic Morphology. In J. Toman (ed.). *Formal Approaches to Slavic Linguistics: The Second Ann Arbor Meeting 2001*. Michigan Slavic Publications: Ann Arbor, MI. 75–96.

Billings, Loren & Catherine Rudin (1996). Optimality and Superiority: A New Approach to Overt Multiple Wh-Ordering. In J. Toman (ed.). *Formal Approaches to Slavic Linguistics: The College Park Meeting 1994*. Michigan Slavic Publications: Ann Arbor, MI. 35–60.

Borsley, Robert D. (1983). A Note on Preposition Stranding. *Linguistic Inquiry* 14(2). 338–343.

Bošković, Željko (1997*a*). On Certain Violations of the Superiority Condition, AgrO, and Economy of Derivation. *Journal of Linguistics* 33(2). 227–254.

Bošković, Željko (1997*b*). Superiority effects with multiple wh-fronting in Serbo-Croatian. *Lingua* 102. 1–20.

Bošković, Željko (1999). On Multiple-Feature Checking: Multiple Wh-Fronting and Multiple Head-Movement. In S. D. Epstein & N. Hornstein (eds). *Working Minimalism*. MIT Press: Cambridge, MA. 159–187.

Bošković, Željko (2001). *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Elsevier Science: Amsterdam.

Bošković, Željko (2002). On Multiple Wh-Fronting. *Linguistic Inquiry* 33(3). 351–384.

Bošković, Željko (2003). On wh-islands and obligatory wh-movement contexts in South Slavic. In C. Boeckx & K. K. Grohmann (eds). *Multiple Wh-Fronting*. John Benjamins: Amsterdam. 27–50.

Bošković, Željko (2005*a*). Left branch extraction, structure of NP, and scrambling. In J. Sabel & M. Saito (eds). *The Free Word Phenomenon: Its Syntactic Sources and Diversity*. de Gruyter: Berlin. 13–73.

Bošković, Željko (2005*b*). On the Locality of Left-Branch Extraction and the Structure of NP. *Studia Linguistica* 59(1). 1–45.

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

Bošković, Željko (2008*a*). On Successive-Cyclic Movement and the Freezing Effect of Feature Checking. In J. Hartmann, V. Hegedus & H. van Riemsdijk (eds). *Sounds of Silence: Empty Elements in Syntax and Phonology*. Elsevier: Amsterdam. 195–233.

Bošković, Željko (2008*b*). On the Operator Freezing Effect. *Natural Language and Linguistic Theory* 26(2). 249–287.

Bošković, Željko (2012). On NPs and clauses. In G. Grewendorf & T. E. Zimmermann (eds). *Discourse and Grammar: From Sentence Types to Lexical Categories*. de Gruyter: Berlin. 179–242.

Bošković, Željko (2016*a*). Getting Really Edgy: On the Edge of the Edge. *Linguistic Inquiry* 47(1). 1–33.

Bošković, Željko (2016*b*). On second position clitics crosslinguistically. In F. Marušič & R. Žaucer (eds). *Formal Studies in Slovenian Syntax: In honor of Janez Orešnik*. John Benjamins: Amsterdam. 23–54.

Chomsky, Noam (1973). Conditions on Transformations. In S. R. Anderson & P. Kiparsky (eds). *A Festschrift for Morris Halle*. Holt, Rinehart and Winston: New York. 232–286.

Chomsky, Noam (1986). *Barriers*. MIT Press: Cambridge, MA.

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

Chomsky, Noam (2000). Minimalist Inquiries: The Framework. In R. Martin, D. Michaels, J. Uriagereka & S. J. Keyser (eds). *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*. MIT Press: Cambridge, MA. 89–155.

Chomsky, Noam (2001). Derivation by Phase. In M. Kenstowicz (ed.). *Ken Hale: A Life in Language*. MIT Press: Cambridge, MA. 1–52.

Cichocki, Wladislaw (1983). Multiple Wh-Questions in Polish. *Toronto Working Papers in Linguistics* 4. 53–71.

Citko, Barbara & Martina Gračanin-Yuksek (2012). Towards a new typology of coordinated wh-questions. *Journal of Linguistics* 49(1). 1–32.

Comorovski, Ileana (1986). Multiple Wh Movement in Romanian. *Linguistic Inquiry* 17(1). 171–177.

Corver, Norbert (1990). *The Syntax of Left-Branch Extractions.* PhD thesis, Tilburg University.

Dayal, Veneeta (1996). *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi.* Kluwer: Dordrecht.

Doron, Edit & Caroline Heycock (1999). Filling and Licensing Multiple Specifiers. In D. Adger, S. Pintzuk, B. Plunkett & G. Tsoulas (eds). *Specifiers: Minimalist Perspectives*. Oxford University Press: Oxford. 69–89.

Elfner, Emily (2016). Stress-Epenthesis Interactions in Harmonic Serialism. In J. J. McCarthy & J. Pater (eds). *Harmonic Grammar and Harmonic Serialism*. Equinox Press: Sheffield. 261–300.

Engdahl, Elisabet (1986). *Constituent Questions: The Syntax and Semantics of Questions with Special Reference to Swedish*. Reidel: Dordrecht.

Fanselow, Gisbert & Anoop Mahajan (2000). Towards a Minimalist Theory of Wh-Expletives, Wh-Copying and Successive-Cyclicity. In U. Lutz, G. Müller & A. von Stechow (eds). *Wh-Scope Marking*. John Benjamins: Amsterdam. 195–230.

Fernández-Salgueiro, Gerardo (2006). On the Incompatibility of Multiple WH-movement with Left-Branch Extraction in Serbo-Croatian. In J. E. Lavine, H. Filip, S. Franks & M. Tasseva-Kurktchieva (eds). *Proceedings of Formal Approaches to Slavic Linguistics 14: The Princeton Meeting 2005*. Michigan Slavic Publications: Ann Arbor, MI. 157–171.

Fox, Danny & David Pesetsky (2005). Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31(1–2). 235–262.

Franks, Steven (2008). Clitic Placement, Prosody and the Bulgarian Verbal Complex. *Journal of Slavic Linguistics* 16(1). 91–137.

Golden, Marija (1997). Multiple wh-questions in Slovene. In W. Browne, E. Dornisch, N. Kondrashova & D. Zec (eds). *Formal Approaches to Slavic Linguistics: The Cornell Meeting 1995*. Michigan Slavic Publications: Ann Arbor, MI. 240–266.

Grebenyova, Lydia (2012). *Syntax, Semantics and Acquisition of Multiple Interrogatives: Who wants what?*. John Benjamins: Amsterdam.

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

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

Haegeman, Liliane, Ángel L. Jiménez-Fernández & Andrew Radford (2014). Deconstructing the Subject Condition in Terms of Cumulative Constraint Violation. *The Linguistic Review* 31(1). 73–150.

Harizanov, Boris (2014). The role of prosody in the linearization of clitics: Evidence from Bulgarian and Macedonian. In C. Chapman, O. Kit & I. Kučerová (eds). *Formal Approaches to Slavic Linguistics 22: The McMaster Meeting 2013*. Michigan Slavic Publications: Ann Arbor, MI. 109–130.

Heck, Fabian & Gereon Müller (2003). Derivational Optimization of Wh-Movement. *Linguistic Analysis* 33(1–2). 97–148.

Heck, Fabian & Gereon Müller (2013). Extremely Local Optimization. In H. Broekhuis & R. Vogel (eds). *Linguistic Derivations and Filtering: Minimalism and Optimality Theory*. Equinox: Sheffield. 136–165.

Heck, Fabian & Gereon Müller (2016). On Accelerating and Decelerating Movement: From Minimalist Preference Principles to Harmonic Serialism. In G. Legendre, M. T. Putnam, H. de Swart & E. Zaroukian (eds). *Optimality-Theoretic Syntax, Semantics and Pragmatics: From Uni- to Bidirectional Optimization*. Oxford University Press: Oxford. 78–110.

Hein, Johannes (2017). *Verbal Fronting: Typology and Theory.* PhD thesis, Universität Leipzig.

Jäger, Gerhard & Anette Rosenbach (2006). The Winner Takes it All – Almost: Cumulativity in Grammatical Variation. *Linguistics* 44(5). 937–971.

Jesney, Karen (2016). Positional Constraints in Optimality Theory and Harmonic Grammar. In J. J. McCarthy & J. Pater (eds). *Harmonic Grammar and Harmonic Serialism*. Equinox: Sheffield. 176–220.

Jurka, Johannes (2010). *The Importance of Being a Complement: CED Effects Revisited.* PhD thesis, University of Maryland.

Kaplan, Aaron (2016). Long-Distance Licensing in Harmonic Grammar. In G. Ó. Hansson, A. Farris-Trimble, K. McMullin & D. Pulleyblank (eds). *Proceedings of the 2015 Annual Meeting on Phonology*. LSA: Washington, D.C.. 1–11.

Kawamura, Tomoko (2004). A feature-checking analysis of Japanese scrambling. *Journal of Linguistics* 40(1). 45–68.

Keine, Stefan (2016). *Probes and their Horizons.* PhD thesis, University of Massachusetts Amherst.

Keller, Frank (2000). *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality.* PhD thesis, University of Edinburgh.

Keller, Frank (2006). Linear Optimality Theory as a Model of Gradience in Grammar. In G. Fanselow, C. Féry, R. Vogel & M. Schlesewsky (eds). *Gradience in Grammar*. Oxford University Press: Oxford. 270–287.

Kimper, Wendell (2011). *Competing Triggers: Transparency and Opacity in Vowel Harmony.* PhD thesis, University of Massachusetts Amherst.

Kimper, Wendell (2016). Positive Constraints and Finite Goodness in Harmonic Serialism. In J. J. McCarthy & J. Pater (eds). *Harmonic Grammar and Harmonic Serialism*. Equinox: Sheffield. 221–235.

Ko, Heejeong (2007). Asymmetries in Scrambling and Cyclic Linearization. *Linguistic Inquiry* 38(1). 49–83.

Ko, Heejeong (2014). *Edges in Syntax: Scrambling and Cyclic Linearization*. Oxford University Press: Oxford.

Koizumi, Masatoshi (1995). *Phrase Structure in Minimalist Syntax.* PhD thesis, MIT.

Kuno, Susumu & Jane J. Robinson (1972). Multiple Wh-Questions. *Linguistic Inquiry* 3(4). 463–487.

Lahne, Antje (2009). A Multiple Specifier Approach to Left-Peripheral Architecture. *Linguistic Analysis* 35(1–4). 73–108.

Larson, Richard K. (1988). On the Double Object Construction. *Linguistic Inquiry* 19(3). 335–391.

Legendre, Géraldine, Antonella Sorace & Paul Smolensky (2006). The Optimality Theory-Harmonic Grammar Connection. In P. Smolensky & G. Legendre (eds). *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Vol. 2. MIT Press: Cambridge, MA. 339–402.

Legendre, Géraldine, Yoshiro Miyata & Paul Smolensky (1990). Can Connectionism Contribute to Syntax? Harmonic Grammar, with an Application. Report CU-CS-485-90, Computer Science Department, University of Colorado at Boulder.

Lipták, Anikó (2009). The Landscape of Correlatives: An Empirical and Analytical Survey. In A. Lipták (ed.). *Correlatives Cross-Linguistically*. John Benjamins: Amsterdam. 1–46.

Mahajan, Anoop K. (2000). Relative Asymmetries and Hindi Correlatives. In A. Alexiadou, A. Meinunger, C. Wilder & P. Law (eds). *The Syntax of Relative Clauses*. John Benjamins: Amsterdam. 201–229.

Manzini, Rita M. (1997). Syntactic Dependencies and their Properties: Weak Islands. In A. Alexiadou & T. A. Hall (eds). *Studies on Universal Grammar and Typological Variation*. John Benjamins: Amsterdam. 135–153.

Matushansky, Ora (2006). Head Movement in Linguistic Theory. *Linguistic Inquiry* 37(1). 69–109.

McCarthy, John J. (2000). Harmonic Serialism and Parallelism. In M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim (eds). *Proceedings of NELS 30*. GLSA Amherst, Massachussets 501–524.

McCarthy, John J. (2008*a*). The Gradual Path to Cluster Simplification. *Phonology* 25(2). 271–319.

McCarthy, John J. (2008*b*). The Serial Interaction of Stress and Syncope. *Natural Language and Linguistic Theory* 26(3). 499–546.

McCarthy, John J. (2010). An Introduction to Harmonic Serialism. *Language and Linguistics Compass* 4(10). 1001–1018.

McCarthy, John J. (2016). The Theory and Practice of Harmonic Serialism. In J. J. McCarthy & J. Pater (eds). *Harmonic Grammar and Harmonic Serialism*. Equinox: Sheffield. 47–87.

Mišmaš, Petra (2015). *On the Optionality of Wh-Fronting in a Multiple Wh-Fronting Language.* PhD thesis, University of Nova Gorica.

Miyagawa, Shigeru (1989). *Structure and Case Marking in Japanese*. Academic Press: San Diego.

Mulders, Iris (1997). Mirrored Specifiers. *Linguistics in the Netherlands* 14. 135–146.

Müller, Gereon (1998). *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Kluwer: Dordrecht.

Müller, Gereon (2001). Order Preservation, Parallel Movement and the Emergence of the Unmarked. In G. Legendre, J. Grimshaw & S. Vikner (eds). *Optimality-Theoretic Syntax*. MIT Press: Cambridge, MA. 278–313.

Müller, Gereon (2011). *Constraints on Displacement: A Phase-Based Approach*. John Benjamins: Amsterdam.

Müller, Gereon (2015). Optimality-Theoretic Syntax. In T. Kiss & A. Alexiadou (eds). *Syntax – Theory and Analysis. An International Handbook*. de Gruyter: Berlin. 875–936.

Müller, Gereon & Wolfgang Sternefeld (2001). The Rise of Competition in Syntax: A Synopsis. In G. Müller & W. Sternefeld (eds). *Competition in Syntax*. de Gruyter: Berlin. 1–68.

Murphy, Andrew (2017). *Cumulativity in Syntactic Derivations.* PhD thesis, Universität Leipzig.

Nichols, Lyn (1999). Movement to Specifiers. In D. Adger, S. Pintzuk, B. Plunkett & G. Tsoulas (eds). *Specifiers: Minimalist Perspectives*. Oxford University Press: Oxford. 206–230.

Nunes, Jairo (2004). *Linearization of Chains and Sideward Movement*. MIT Press: Cambridge.

Pater, Joe (2009). Weighted Constraints in Generative Linguistics. *Cognitive Science* 33. 999–1035.

Pater, Joe (2012). Serial Harmonic Grammar and Berber syllabification. In T. Borowsky, S. Kawahara, T. Shinya & M. Sugahara (eds). *Prosody Matters: Essays in Honor of Elisabeth O. Selkirk*. Equinox: London. 43–72.

Pater, Joe (2014). Candian Raising with Language-Specific Weighted Constraints. *Language* 90(1). 230–240.

Pater, Joe (2016). Universal Grammar with Weighted Constraints. In J. J. McCarthy & J. Pater (eds). *Harmonic Grammar and Harmonic Serialism*. Equinox: Sheffield. 1–46.

Polinsky, Maria, Carlos G. Gallo, Peter Graff, Ekaterina Kravtchenko, Adam Milton Morgan, & Anne Sturgeon (2013). Subject Islands are Different. In J. Sprouse & N. Hornstein (eds). *Experimental Syntax and Island Effects*. Cambridge University Press: Cambridge. 286–309.

Postal, Paul M. (1998). *Three Investigations of Extraction*. MIT Press: Cambridge, MA.

Potts, Christopher, Joe Pater, Karen Jesney, Rajesh Bhatt & Michael Becker (2010). Harmonic Grammar with Linear Programming: From Linear Systems to Linguistic Typology. *Phonology* 27(1). 77–117.

Prince, Alan & Paul Smolensky (1993/2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Wiley-Blackwell: Oxford.

Rezac, Milan (2004). *Elements of Cyclic Syntax: Agree and Merge*. PhD thesis, University of Toronto.

Richards, Norvin (1999). Featural cyclicity and the ordering of multiple specifiers. In S. D. Epstein & N. Hornstein (eds). *Working Minimalism*. MIT Press: Cambridge, MA. 127–158.

Richards, Norvin (2001). *Movement in Language: Interactions and Architectures*. Oxford University Press: Oxford.

Richards, Norvin (2016). *Contiguity Theory*. MIT Press: Cambridge, MA.

Rizzi, Luigi (1996). Residual verb second and the *Wh*-Criterion. In A. Belletti & L. Rizzi (eds). *Parameters and Functional Heads: Essays in Comparative Syntax*. Vol. 2. Oxford University Press: Oxford. 63–90.

Ross, John R. (1967). *Constraints on Variables in Syntax*. PhD thesis, MIT.

Ross, John R. (1986). *Infinite Syntax!*. Ablex Publishing: Norwood.

Ross, John R. (1987). Islands and Syntactic Prototypes. In A. Bosch, B. Need & E. Schiller (eds). *Papers from the Twenty-Third Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society: Chicago, IL. 309–320.

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

Ryan, Kevin (2017). Attenuated Spreading in Sanskrit Retroflex Harmony. *Linguistic Inquiry* 48(2). 299–340.

Sabel, Joachim (2001). Deriving Multiple Head and Phrasal Movement: The Cluster Hypothesis. *Linguistic Inquiry* 32(3). 532–547.

Sabel, Joachim (2005). String-vacuous Scrambling and the Effect on Output Condition. In J. Sabel & M. Saito (eds). *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity*. de Gruyter: Berlin. 281–333.

Saito, Mamoru (1985). *Some Asymmetries in Japanese and their Theoretical Implications.* PhD thesis, MIT.

Scott, Tatiana V. (2012). *Whoever doesn't HOP must be Superior: The Russian Left-Periphery and the Emergence of Superiority.* PhD thesis, Stony Brook University.

Speas, Margaret (1990). *Phrase Structure in Natural Language*. Springer: Dordrecht.

Srivastav, Veneeta (1991). The Syntax and Semantics of Correlatives. *Natural Language and Linguistic Theory* 9(4). 637–686.

Stjepanović, Sandra (1999). *What do second position cliticization, Scrambling and multiple wh-fronting have in common?.* PhD thesis, University of Connecticut.

Stjepanović, Sandra (2003). Multiple Wh-Fronting in Serbo-Croatian Matrix Questions and the Matrix Sluicing Construction. In C. Boeckx & K. K. Grohmann (eds). *Multiple Wh-Fronting*. John Benjamins: Amsterdam. 255–284.

Takano, Yuji (2002). Surprising Constituents. *Journal of East Asian Linguistics* 11(3). 243–301.

Toman, Jindřich (1981). Aspects of Multiple Wh-Movement in Polish and Czech. In R. May & J. Koster (eds). *Levels of Syntactic Representation*. Foris: Dordrecht. 293–302.

Torres-Tamarit, Francesc (2012). *Syllabification and Opacity in Harmonic Serialism.* PhD thesis, Universitat Autònoma de Barcelona.

Travis, Lisa (1984). *Parameters and Effects of Word Order Variation.* PhD thesis, MIT.

Ura, Hiroyuki (1996). *Multiple Feature-Checking: A Theory of Grammatical Function Splitting.* PhD thesis, MIT.

van Urk, Coppe & Norvin Richards (2015). Two Components of Long-Distance Extraction: Successive Cyclicity in Dinka. *Linguistic Inquiry* 46(1). 113–155.

Vicente, Luis (2009). An Alternative to Remnant Movement for Partial Predicate Fronting. *Syntax* 12(2). 158–191.

Wachowicz, Krystyna A. (1974). Against the Universality of a Single Wh-Question Movement. *Foundations of Language* 11(2). 155–166.

Wilder, Chris (1997). English Finite Auxiliaries in Syntax and Phonology. In J. R. Black & V. Motapanyane (eds). *Clitics, Pronouns and Movement*. John Benjamins: Amsterdam. 321–362.

Williams, Edwin (2003). *Representation Theory*. MIT Press: Cambridge, MIT.

Zwart, Jan-Wouter (1997). Transitive Expletive Constructions and the Evidence Supporting the Multiple Specifier Hypothesis. In W. Abraham & E. van Gelderen (eds). *German: Syntactic Problems – Problematic Syntax*. Niemeyer: Tübingen. 105–134.