1996 - FASL 3

Optimality and Superiority:
A new approach to overt multiple-\textit{wh} ordering

Loren Billings, Florida State University
Catherine Rudin, Wayne State College

In this paper we account for Superiority effects in Bulgarian using mechanisms of Optimality Theory (Prince & Smolensky 1993), as applied to syntax by Grimshaw (1993a) and specifically to Superiority in Grimshaw (1993b). Our primary theoretical innovations are constraints to account for animacy and for certain surface effects of consecutive homophonous \textit{wh} words, as well as constraints that serve to distinguish languages like Bulgarian—which front an indefinite number of \textit{wh} phrases in overt syntax—from those that front one (as does English) or none at all (Chinese, for example).

1 Summary of the problem

Wachowicz (1974) introduced the generative enterprise to the violability of single \textit{wh}-fronting as a universal, using Polish data.\textsuperscript{1} Rudin (1988a; 1988b; 1989) shows that Bulgarian is different from the other Slavic languages, but like Romanian (and possibly Romany and Yiddish) in fronting all \textit{wh} phrases to SpecCP in overt syntax; cf. Comorovski (1986; 1989) on Romanian and Lakova (1991) on Bulgarian. The other Slavic languages, Rudin (1988b) adds, front only once to SpecCP; any other \textit{wh} phrases are adjoined to IP—essentially equivalent to Toman’s (1981) pre-CP model.

1.1 Evidence for a single syntactic \textit{wh} constituent in Bulgarian

Rudin’s multiply-filled-SpecCP model is supported by several empirical distinctions between Bulgarian on the one hand, and the rest of Slavic on the other. Two such distinctions are shown here, with the Serbo-Croatian examples in (2) and (4) serving as representative of the other Slavic languages. While either order is allowed in (2), only one order is allowed in (1)—in which the \textit{wh} phrases form an uninterrupted unit. Rudin (1988b) also shows that Bulgarian parentheticals behave similarly, in that they must follow all fronted \textit{wh} phrases.\textsuperscript{2}
Bulgarian

(1a) Zavisí ot tova, …
   koj kogo práv e udaril.
   who whom first CL hit
   NOM ACC ADV 3.SG M.SG

(1b) Zavisí ot tova, …
   *koj práv kogo e udaril.
[= ex. (42), Rudin (1988b:467)]

Serbo-Croatian

(2a) Zavisí od toga …
   ko koga prv i udari?
   who whom first hit
   NOM ACC ADV M.SG

(2b) Zavisí od toga …
   ko prv koga udari?
[= ex. (10a-b), Bošković (1994:6)]

‘It depends who hit whom first.’ (same gloss for both)

Additionally, Bulgarian requires all wh phrases to extract to a higher (non-wh) clause, as shown in (3). Extraction is required in (4), but only once—without specifying which wh phrase. Only Bulgarian requires all fronted wh phrases to appear together.

Bulgarian

(3a) Koj kâde misliš [če e otišāl]?
   who where think gone
   NOM ADV 2.SG C CL M.SG

(3b) *Koj misliš [če kâde e otišāl]?
[= ex. (13), Rudin (1989:6)]

(3c) *Kâde misliš [če koj e otišāl]?

Serbo-Croatian

(4a) Ko šta želite [da vam kupi]?
   who what want you buy
   NOM ACC 2.PL C DAT 3.SG

(4b) Ko šta želite [da vam kupi]?

(4c) Šta želite [da vam ko kupi]?

‘Who do you think went where?’ ‘Who do you want to buy what?’

1.2 The Superiority Condition, as applied to Bulgarian

The Superiority Condition,3 proposed in Chomsky (1973), as applied to wh-movement, is the constraint in languages like English that assures that the syntactic subject and not the object is fronted if both are wh phrases. Rudin (1985; 1986; 1988a; 1988b; 1989) shows that there is a Superiority-like effect in Bulgarian, requiring specific ordering among multiply fronted wh phrases. That is, a fronted subject wh phrase must precede a fronted non-subject wh phrase (Rudin 1985:2; 1986:120, fn. 40). This is shown, for example, in (1a) or (3a) above. Rudin (1986:118) does, however,
mention that this generalization is “not entirely accurate”, offering the “rules of thumb” in (5), adding that (5a) and (5b) are exceptionless, but (5d) is not, as is evidenced by (6). As (7) shows, however, (5d) is strong enough to override (5e).

(5) **Rules of thumb for wh ordering in Bulgarian**
   a. NOM **koj** ‘who’ is always first;
   b. a **wh word** must precede a **wh prepositional phrase** containing the same **wh word** (including DAT **na kogo**);
   c. all else being equal, a human **wh word** precedes a non-human one;
   d. NOM/ACC **kakvo** ‘what’ tends to be second; and
   e. **wh adverbials** tend to be late in the series of **wh phrases**.

(6) **Koj** na **kogo** **kakvo** e kazal?
   who NOM to whom DAT what ACC CL 3.SG said M.SG
   ‘Who said what to whom?’

(7) **Koga** **kakvo** e kupil?
   when what ACC CL 3.SG bought M.SG
   ‘When did he buy what?’

In this paper we show that not only are (5d) and (5e) violable, so is (5a). Additionally, (5c) is, as worded, also violable. A more explicit system of constraints is presented below (in §5), allowing optional **wh** orders in some environments and rigid ones in others. First, however, an overview of the data (in §2–§3), followed by some new data from colloquial Bulgarian (in §4).

2 The simple clause with two **wh** phrases: **NOM + ACC**

The clearest data on the Superiority Condition have been clauses with a NOM **wh** external argument and an ACC **wh** direct object. Some **wh** questions in Bulgarian do apparently violate Superiority.

Three of the rules of thumb in (5) rely on the notion of animacy: (5c) directly, while (5a) and (5d) do so implicitly. The
data below show that (5a) is unviolated so long as koi is the external argument (or the only human argument; we return to such distinctions in §3.1). Suffice it to say that in transitive sentences koi ‘who’ is always first in the wh cluster, regardless of the wh direct object’s animacy, as shown in examples (8) and (9):

(8a)  
\[ \text{Ko}\text{j \ kogo } \text{vižda?} \]
\[ \text{who NOM whom ACC sees 3.SG} \]
\[ '\text{Who sees whom.'} \]
\[ [\text{~ ex. (55a-b), Rudin (1988b:473)}] \]

(8b)  
\[ \text{*Kogo koj vižda?} \]
\[ \text{who NOM whom ACC sees 3.SG} \]

(9a)  
\[ \text{Ko}\text{j \ kakvo } \text{pravi?} \]
\[ \text{who NOM what ACC does 3.SG} \]
\[ '\text{Who does what.'} \]
\[ [\text{~ ex. (75a-b), Rudin (1988b:481-2)}] \]

The datum heretofore missing from the literature, to our knowledge, is the multiple question with an inanimate wh external argument and a human wh internal argument. An agentive transitive predicate like hit shows this most effectively. In English such questions must conform to the Superiority Condition, as shown:

(10a)  
\[ \text{What hit whom?} \]

(10b)  
\[ \text{*Whom (did) what hit?} \]

(11a)  
\[ \text{I know what hit whom.} \]

(10b)  
\[ \text{*I know whom what hit.} \]

Regardless of embedding, as is the case in (11), the (a) examples are the only grammatical means of expressing such questions.  

In Bulgarian either both orders in (12) are acceptable, or—

— with certain informants—one or the other is more natural. Our informants all find both (12a) and (12b) minimally acceptable.  

(12a)  
\[ \text{Ko}\text{go \ kakvo } \text{e udarilo?} \]
\[ \text{whom ACC what NOM CL 3.SG hit N.SG} \]
\[ '\text{What hit whom?'} \]

(12b)  
\[ \text{Kakvo \ kogo } \text{e udarilo?} \]
\[ \text{what NOM whom ACC CL 3.SG hit N.SG} \]

Embedding the wh clause within another clause (not shown here) does not appear to affect our informants' judgments. It is also
worth mentioning that informants tend to try a number of other truth-value equivalents of (12), primarily by trying to passivize the question. This makes sense, considering the general tendency in Slavic for the NOM external argument of a transitive verb to be interpreted as an Agent, as well as another tendency—across human language—for an inanimate subject not to be an Agent.

While the data are somewhat murky, a few conclusions can be drawn: First, it looks as if a purely syntactic Superiority account does not receive any support here. Next, it is likely that our informants had never encountered data like (12), in which animacy and Superiority are teased apart. Regardless of which view one takes about how parameters are set, it can safely be said that animacy and Agent-hood may never have been differentiated during acquisition with positive evidence. We return to this issue (in §5.2) below.

3 Other syntactic combinations of wh phrases

3.1 NOM + DAT
We distinguish between two types of DAT arguments: expressions of Goal and Experiencer theta (= thematic, = semantic) roles:

3.1.1 NOM Agent + DAT Goal: Not surprisingly, a NOM external argument must precede a DAT internal argument when both are wh expressions, as is shown in (13):

(13a)  \[ \text{Koj} \quad \text{na kogo e} \quad \text{dal} \quad \text{ximikalkata?} \]
who NOM to whom DAT CL.3.SG gave M.SG pen-the ACC

(13b)  \[ \text{Na kogo koj e} \quad \text{dal} \quad \text{ximikalkata?} \]
\[ \text{Who gave the pen to whom?} \]  \[ (= \text{ex. (6), Rudin (1985:2))} \]

3.1.2 NOM Theme + DAT Experiencer: There are also constructions in which the DAT expresses the Experiencer theta role while the NOM expresses the Theme role. Example (14) shows that two human wh phrases with these respective cases and roles can have either order. This theta-to-case permutation cannot be tested with a non-human internal argument, because such dative arguments are recipients, which are, at the very least, animate, requiring the use of \( \text{na kogo} \) ‘to whom’ (instead of some case form of the kakvo stem).
(14a) \textit{Koj na kogo mu xaresva?}  
who NOM to whom DAT CL.DAT.3.SG is-pleasing 3.SG

(14b) \textit{Na kogo koj mu xaresva?}  
(literally) 'Who is likable to whom?'

Both forms in (14) are acceptable, though some informants favor one or the other. Example (15) has non-human \textit{kakvo} as Theme:

(15a) ?? \textit{Kakvo na kogo mu xaresva?}  
what NOM to whom DAT CL.DAT.3.SG is-pleasing 3.SG

(15b) \textit{Na kogo kakvo mu xaresva?}  
(literally) 'What is likable to whom?'

Comparing (15) to (14) reveals that a human DAT Experiencer \textit{wh} phrase is not obligatorily ordered with respect to a \textit{wh} human NOM Theme, but should be first when the \textit{wh} Theme is \textbf{non-human}.

In (15) the strong preference is for the human (Experiencer) \textit{wh} phrase to be uttered first. These examples, in addition to the \textit{what-hit-whom} examples in (12), suggest the following: First, Superiority holds if the NOM \textit{wh} phrase is both the external argument and human. If a \textit{wh} external argument is non-human or if a human NOM \textit{wh} phrase is not the external argument, then another \textit{wh} word can optionally appear first. When the NOM \textit{wh} phrase is neither human nor the external argument, and there is a human \textit{wh} Experiencer, then the strong preference is for the NOM \textit{wh} phrase not to be initial. Our constraints below (in §5.1) capture these seemingly disjoint generalizations.

3.2 \textit{ACC direct object + DAT indirect object}
Especially interesting in the recent literature are the proposed structures of the DAT and ACC internal arguments of a verb. Certain recent proposals, most notably in Bailyn (1995), argue that the DAT indirect object (IO) is the complement of V while the ACC DO is generated in SpecVP. Our data in this area do not add to the picture \textit{per se}. But we do add one crucial clarification to one apparent DAT-ACC asymmetry in the Superiority literature on Bulgarian. Much of the work on Superiority accounts for the apparent lack of ordering
of DO and IO *wh* phrases using the generative notions "minimal domain" and "*m*-command", which both essentially render the DO and IO positions equidistant to their common SpecCP destination. The examples in (16) show a three-place predicate with all of its arguments human, and only the DO and the IO are *wh* phrases.

(16a)  
\[ \textit{Kogo na kogo e pokazal Ivan?} \]  
whom ACC to whom DAT CL 3.SG showed M.SG Ivan NOM

(16b)  
\[ \textit{Na kogo kogo e pokazal Ivan?} \]  
‘Whom did Ivan show to whom?’  
[= ex. (5), Rudin (1985:2)]

It looks as though the DO must be superior to (or higher in the syntactic tree than) the IO. We return to this issue (in §4) using data from colloquial Bulgarian, showing that there is no syntactic *wh*-ordering requirement in (16), but merely a surface constraint, and this factor can be conveniently controlled for in colloquial Bulgarian.

3.3 The order of *wh* arguments and *wh* adverbials

Rudin's last (1986) rule of thumb ("adverbs tend to be late in the series of *wh* phrases," = (5e) above) is just that: a tendency. The placement of *wh* adverbials requires only a minor amendment to the description so far: Syntactically speaking, *wh* adverbs behave identically to *kogo 'whom' (ACC) or *kakvo 'what' (NOM or ACC). That is, NOM *koj 'who' precedes a *wh* adverb, whereas *kogo 'whom' and *kakvo 'what' (regardless of grammatical relation) are not ordered relative to *wh* adverbs.

Consistent with the data above, NOM *koj 'who' must precede a *wh* adverbial:

(17a)  
\[ \textit{Koj kak pătuva?} \]  
who NOM how travels 3.SG (i.e., by what conveyance)

(17b)  
\[ \textit{*Kak koj pătuva?} \]

(18a)  
\[ \textit{Koj kăde Šte spi?} \]  
who NOM where MODAL sleeps

(18b)  
\[ \textit{*Kăde koj Šte spi?} \]
(19a) \( Koj \ zaštto \ ti \ xaresva? \)
\[ \text{who NOM why CL DAT.2.SG is-pleasing 3.SG} \]
(19b) * \( Zaštto \ koj \ ti \ xaresva? \)
'Who do you like why?'

The ACC wh word kogo 'whom' appears in either order with a wh adverbial. Bošković (1994 and p.c.) reports that whereas the order in (20a) is always acceptable, his informants judge the other, adverbial-initial order less than perfect, ranging from slightly bad (?) to near ungrammaticality (*?), often depending on exactly which wh adverbial is used. We have elicited judgments. We leave this issue open for future research. We might add only that our informants were usually able to accept orders such as those in (20b), but only after conceptualizing the necessary, non-neutral context. We do not consider questions with multiple wh-adverbials.

(20a) \( Kogo \ kāde \ ste \ videli? \)
\[ \text{whom ACC where CL 2.PL saw PL} \]
(20b) \( Kāde \ kogo \ ste \ videli? \)
'Whom did you see where?'

Regardless of case, kakvo 'what' and wh adverbials are not ordered:

(21a) \( Kakvo \ koga \ e \ kupil? \)
\[ \text{what ACC when CL 3.SG bought M.SG} \]
(21b) [= (7)] \( Koga \ kakvo \ e \ kupil? \)
'When did he buy what?'

(22a) \( Kakvo \ kāde \ raste? \)
\[ \text{what NOM where grows 3.SG} \]
(22b) \( Kāde \ kakvo \ raste? \)
'What grows where?'

To summarize the wh adverbials, the only requirement on their order—with regard to wh arguments—is that NOM koj 'who' must precede them. Note that koj is not always required to appear first (cf. (14b) above). The relative ordering of a wh adverbial is
free so long as the only other wh phrases are ACC kogo ‘whom’, DAT na kogo ‘to whom’, or NOM/ACC kakvo ‘what’.

3.4 Ordering of subsequent (non-initial) wh constituents
We turn now to whether the second and third (or more) wh phrases are ordered. We repeat some of the polemics on this issue here briefly: According to Rudin (1988b:472), a NOM wh word must precede an ACC wh word and “when a Wh-word indirect object is also present, the order of the three Wh-words must be subject, direct object, indirect object” [p. 472], providing the example in (23a).

(23a) Koj kogo na kogo e pokazal?
who NOM whom ACC to whom DAT CL 3.SG showed M.SG
(23b) *Koj na kogo kogo e pokazal?
‘Who showed whom to whom?’ [= (54c), Rudin (1988b:473)]

Dimitrova-Vulchanova (1992:45) responds that “the order of DO and IO Wh-constituents is not fixed”, providing the following:

(24a) Koj kakvo na kogo kaza?
who NOM what ACC to whom DAT said 3.SG
(24b) Koj na kogo kakvo kaza?
‘Who said what to whom?’
[= ex. (80a-b), Dimitrova-Vulchanova (1992:45)]

Dimitrova-Vulchanova is right. There is no a priori precedence restriction between direct and indirect objects in Bulgarian.

Nevertheless, the examples in (23)—and, for that matter, (16)—are restricted to a single ordering of wh elements. We offer a non-syntactic, non-discourse, non-functionalist explanation of the distribution in (23) and (24). Namely, a “low-level, PF-leg, surface-output” constraint. Note that (16b) and (23b) both have the sequence na kogo kogo (literally ‘to whom whom’). Recall Rudin’s (1985:118) second rule of thumb: “A wh word must precede a wh prepositional phrase containing the same wh word (including DAT na kogo)” [= (5b)]. We expound on this observation in the next section using additional colloquial data.
4 Comparison with the colloquial register

So far we have shown that *na kogo kogo ‘to whom whom’ is ruled out by consecutive wh homophones. Rudin (1986:9) briefly mentions that kogo ‘whom’ is sometimes replaced by koj ‘who’ in colloquial speech, similar to the replacement of whom with who in English. Judgments of colloquial data vary, but for some of our informants the DAT is expressed as na koj, literally ‘to who’, while the ACC often remains as kogo ‘whom’, especially when NOM koj ‘who’ also appears in the wh cluster or if the sentence is pragmatically strange, as in the what-hit-whom examples in (12). The wh vocabulary of this colloquial register is outlined in (25):

(25) Colloquial Bulgarian:

a. NOM  b. ACC  c. DAT

koj  kogo  or  koj  na koj

‘who’  ‘whom’  ‘who’  ‘to who’

This wh sublexicon has a way of avoiding successive instances of kogo ‘whom’, but does cause the DAT-NOM sequence na koj koj (literally ‘to who who’). Both (literary) na kogo kogo and (colloquial) na koj koj are ruled out. We show this by repeating any of the data with a differing grammaticality judgment when na kogo ‘to whom’ is replaced by na koj ‘to who’. (In extremely colloquial examples like Koj koj trjabva da sluša? ‘Who has to obey who?’, where koj is also ACC, the two instances of koj are forced together. Such examples are apparently preferable to leaving one in situ.)

Literary Bulgarian (repeated)

(14a)  Koj  na kogo  mu  xaresva?
(14b)  Na kogo  koj  mu  xaresva?

Colloquial Bulgarian

(26a)  Koj  na koj  mu  xaresva?
        who NOM  to who DAT  CL.DAT.3.SG  is-pleasing 3.SG

(26b)  *Na koj  koj  mu  xaresva?

(literally) ‘Who is pleasing to who?’
Compare also the counterparts in (16) and (27); their (b) examples show that there is no **syntactic** superiority at play here.

**Literary Bulgarian** (repeated)

(16a) *Kogo na kogo e pokazal Ivan?*

(16b) *Na kogo kogo e pokazal Ivan?*

**Colloquial Bulgarian**

(27a) *Kogo na koi e pokazal Ivan? whom ACC to who DAT CL.3.SG showed M.SG Ivan NOM*

(27b) *Na koi kogo e pokazal Ivan?*

(literally) ‘Whom did Ivan point out to who?’

Finally, compare the counterparts in (23) and (28):

**Literary Bulgarian** (repeated)

(23a) *Koj kogo na kogo e pokazal?*

(23b) *Koj na kogo kogo e pokazal?*

**Colloquial Bulgarian**

(28a) *Koj kogo na koi e pokazal? whom NOM whom ACC to who DAT CL.3.SG showed M.SG*

(28b) *Koj na koi kogo e pokazal?*

(literally) ‘Who pointed out whom to who?’

Clearly the contrasts between these two registers of Bulgarian suggest that syntax is not involved ruling out (16b) or (23b)—or, for that matter, the colloquial example in (26b).

Some sort of constraint is required to rule out sequences such as *na kogo kogo* and *na koj koj*. In none of the examples so far is this constraint violated. This type of constraint is not unheard-of in other languages. Kornfilt (1986) proposes the “Stuttering Prohibition” to rule out consecutive sequences of a compound marker and possessive marker in Turkish, which can be homophonous. Ross (1972) proposes a “Doubl-Ing” surface
constraint for English. Napoli (1976) describes how two clitics in Italian, usually both pronounced [si], when together are pronounced [či si]. There is also a similar sort of effect in Russian and Polish nominalizations, where there is a strong tendency against two arguments with the same case (see, for example, Dziworek 1993, Rappaport 1992). We call this constraint STARHOM.9

A final note on the colloquial examples in this section: In all the sets compared here—(14) with (26), (16) with (27), and (23) with (28)—one register has an optional order while the other has only one allowed order. That is, all of the example pairs have syntactically unordered wh clusters that are further restricted as in one or the other register due to consecutive homophony. We have not shown, therefore, that consecutive homophony can override Superiority, only that it can further restrict the set of grammatical outputs of a different grammar component.

5 Formal definitions, constraints proposed

Prior to proposing any more constraints we summarize the data:

- A wh external argument has at least the option of appearing first:
  — If the wh external argument is non-human (i.e., kakvo), and there is a human wh internal argument, then the wh-ordering is optional (as in example (12) above).10
  — If the wh external argument is human (i.e., koj), then it must appear first in the wh cluster (cf. (8), (9), (13) or (17)11).

- If the external argument is not a wh phrase, as in (16) and (20)-(21), then there can be any ordering in the wh cluster.

- If there is no external argument of any kind (regardless of whether the external argument is a wh phrase), and:
  — if two (or more) wh phrases are human, as in (14), then any of the human wh phrases can be first in the wh cluster.
  — if both (or all) wh phrases are non-human, as in (22), then any of these can be first in the wh cluster.
  — if there is only one human wh phrase, as in (15) and (18)-(19), then that wh phrase must appear first in the wh cluster.
5.1 The notions "sorting key" and "subject"

One proposal in the literature is that the first wh phrase in a clause is a "sorting key"; we interpret this to mean formally that the sorting-key wh phrase represents the left-hand column of items in a logical function (i.e., the column the items of which may not recur).

(29) **Kuno's Sorting Key Hypothesis**

"In a multiple wh question, the leftmost wh-word represents the key for sorting relevant pieces of information in the answer." [Kuno & Takami (1993:112), citing Kuno (1982)]

Indeed, the Bulgarian data above show that wherever there is a syntactic ordering requirement (that is, ignoring consecutive homophony) a wh phrase may be required to be only first in a cluster (i.e., never required to be second, third, or last in the cluster). Whereas some of the Rules of Thumb in (5) describe how some wh phrases tend to be second or late in the series, none of them absolutely requires such non-initial orderings.

It is also worth pointing out that in a single-wh-fronting language like English a wh phrase in situ must be dependent on a preceding (fronted) wh phrase. This means that a wh clause licenses only one wh phrase, the sorting key. Any others are dependent on the sorting key. We adopt the approach to Superiority in Williams (1994:191ff), which relies on time-line precedence. The only difference between English and Bulgarian is that dependent wh phrases must also front in overt syntax. In both languages such dependent wh phrases must still follow (i.e., appear to the right of) the sorting key. (We return to the formal constraints distinguishing these two language types below.) We adopt the Sorting Key Hypothesis and Williams's proposal that wh dependence is subsumed in the Leftness Condition, and that they are unviolated (i.e., part of “Gen” in Prince & Smolensky’s 1993 model).

A brief excursus on what the sentential subject is will also clarify the proposals below: In the talk version of this paper at FASL-3 we assumed that the clausal subject is whichever NP is in the NOM case. Thanks to a suggestion by Željko Bošković, we now assume otherwise (although perhaps not in keeping with his exact intention): NOM is assigned by default; DAT and ACC are assigned
to specific internal arguments by virtue either of their position within VP (or, possibly in the case of the DAT, idiosyncratically in the lexicon). This entails that any external argument (in situ) bears NOM case. So does any internal-argument Theme not already assigned ACC. The canonical subject position, SpecIP, no longer functions just as a Case-assigning position, but as a marker of some sort of prominence, perhaps following Izvorski (1993). The structural positions of crucial significance are the following:

\[(30) \quad [CP \ [IP \ X \ [PrP \ X \ [VP \ ... \ X^n \ ... ]VP \ ]PrP \ ]IP \ ]CP\]

Pr[edicate]P is a “VP shell”, a projection that generates an external argument if there is one; cf. Baily (1995) and Bowers (1993) for further details. Each instance of X in (30) represents a position (or, inside VP, positions) where wh phrases are located (prior to wh-movement). An external argument is projected in SpecPrP, while adverbials and internal arguments are within VP. While we do not make the crucial arguments for the VP-internal status of adverbs here (or, for that matter, specify where within VP they are positioned), our analysis functions properly under this (albeit vague) assumption. We posit no particular order within VP in our structures. Furthermore, SpecIP is filled as the result of a pre-wh, Move-Alpha operation. We propose the following constraints:

\[(31) \quad \text{SUBJSUP: Fill SpecIP with the highest XP within IP.}\]

If there is a PrP, then the highest XP is SpecPrP, the external argument. If there is no PrP, then all constituents are within VP and none of these is higher than any other; any VP-internal constituent can be moved to SpecIP in such a structure to satisfy SUBJSUP.

\[(32) \quad \text{SUBJHUM: SpecIP must be human.}\]

SUBJHUM is a constraint that formalizes the interaction of pragmatics with syntax; Optimality Theory is especially suited to such interface constraints. Neither of the constraints in (31) and (32) refers to wh-hood. A non-wh external argument in SpecIP, for example, will satisfy both of these constraints.
We further assume a canonical (structural) version of the Superiority Condition, worded here in terms of (29) above:

(33) **SORTSUP**: The structurally highest *wh* phrase (in an A[rgument] position) must be the sorting key.

**SORTSUP** merely insures that the *wh* phrase arboreally highest (prior to any *wh*-movement) appear first (leftmost) in overt syntax. The interaction of these three constraints is exemplified in §5.2 below. By "*wh*-movement" we do not mean "movement to SpecCP"; we follow Grimshaw (1993a, b), which argues (using data primarily from English) that there is no separate CP projection if SpccIP is occupied by a *wh* phrase.14

### 5.2 The tableaux

A representative of each of the structure types summarized at the start of §5 is assessed using Optimality-theoretic tableaux. The three constraints in (31) through (33) correspond to columns on the right side of each tableau. (STARHOM, from §4, is not shown, since it is unviolated in all the remaining data; cf., however, n. 16 below) Likely candidates are arranged in rows. A star (*) in any cell signifies that the candidate form fails that constraint column. A check mark (✓) signifies that there is no violation; a dollar sign ($) marks each tableau’s optimal candidate(s).

<table>
<thead>
<tr>
<th>(34) [= (8)]</th>
<th>SORT</th>
<th>SUP</th>
<th>SUBJ</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. [ikkoji [P,P NP,Pi] [VP...kogo join...]]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>$ [ikkoji kogo join [P,P NP,Pi] [VP...wh join...]]</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>b.I [ikkogo join [P,P koji [P,P wh join [VP...NP join...]]]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>b.II [ikkoji [P,P NP,Pi] [VP...kogo join...]]</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ikkogo join [P,P koji [P,P wh join [VP...NP join...]]]</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three candidates are displayed in (34). The lower two candidates correspond to possible structures underlying the bad word order in (8b). For ease of exposition, we have listed each candidate in (34)
in two tiers: The top tier is the structure “before” wh-movement, while the lower tier reflects the structure at overt syntax.

Candidate (34a) satisfies all three constraints: koj ‘who’ is human and base-generated higher than kogo ‘whom’. As such, this wh phrase satisfies both SUBJHUM and SUBJSUP. Additionally, since it is in SpecIP (i.e., the arboreally highest wh phrase in the clause) koj also satisfies SORTSUP. Candidate (34b.I) fails SUBJSUP because the lower argument is NP-moved to SpecIP. Candidate (34b.II) satisfies SUBJHUM and SUBJSUP, but—due to the precedence of kogo ‘whom’—causes the (pre-wh) lower wh phrase to be interpreted as the sorting key, in violation of SORTSUP. (In the remaining tableaux we show only the input to wh-movement. All of the candidates we show go on to undergo wh-movement in accordance with SORTSUP. For this reason, only SUBJSUP and SUBJHUM columns appear in the remaining tableaux.)

Tableau (35) is similar to (34). The only difference is that candidate (35b) also fails SUBJHUM.

<table>
<thead>
<tr>
<th>(35) [\approx (11)]</th>
<th>SUBJ</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $ [[pkoji] [p_p NP_i [vp...kakvoj...]]]</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>b. $ [[pkakvoj [p_p koji [vp...NP_i...]]]</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

Tableau (36) presents several possibilities: It may well be that the average Bulgarian-speaker has never encountered the positive evidence to rank SUBJSUP and SUBJHUM. If a speaker has not encountered such (pragmatically odd) data, then there is no way for these constraints to be ranked. It is also possible that other focus-type constraints are at play. We leave open the implications.

<table>
<thead>
<tr>
<th>(36) [\approx (12)]</th>
<th>SUBJ</th>
<th>SUBJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. $ [[pkakvoi [p_p NP_i [vp...kogoj...]]]</td>
<td>√</td>
<td>*</td>
</tr>
<tr>
<td>b. $ [[p_kogoj [p_p kakvoi [vp...NP_i...]]]</td>
<td>*</td>
<td>√</td>
</tr>
</tbody>
</table>
Tableau (37) has a non-wh human external argument NP-moved to SpecIP. The forms in (37) are the “input” to wh-movement. While NP-moving any of the arguments to SpecIP satisfies SUBJ_HUM, only Ivan can be so moved to satisfy SUBJ_SUP.

\[(37) \equiv (16)\]

\[
\begin{array}{c|c|c|c}
 & \text{SUBJ SUP} & \text{SUBJ HUM} \\
 a. & [\text{IP Ivan} \ [\text{P NP} \ [\text{VP...kogo...na kogo...]...]]] & \checkmark & \checkmark \\
b. & [\text{IP kogo} \ [\text{P NP} \ [\text{VP...NP} \ [\text{VP...na kogo...]]...]]] & * & \checkmark \\
c. & [\text{IP na kogo} \ [\text{P NP} \ [\text{VP...kogo...NP} \ [\text{VP...NA...]]...]]] & * & \checkmark \\
\end{array}
\]

Because Ivan satisfies both of the NP-movement constraints, either wh phrase is then free to be sorting key. The structures in (38a-b) correspond to the two optional results of wh-movement applied to (37a). Both forms in (38), in turn, satisfy SORTSUP.16

\[(38a) \quad [\text{CP kogo na kogo} \ [\text{IP Ivan} \ [\text{P NP} \ [\text{VP...whj...whk...]]...]]] \\
(38b) \quad [\text{CP na kogo kogo} \ [\text{IP Ivan} \ [\text{P NP} \ [\text{VP...whj...whk...]]...]]]
\]

The remaining tableaux assess the structures with no external argument of any kind (i.e., without any Pr projection).

\[(39) \equiv (14)\]

\[
\begin{array}{c|c|c|c}
 & \text{SUBJ SUP} & \text{SUBJ HUM} \\
 a. & \checkmark & \checkmark \\
b. & \checkmark & \checkmark \\
\end{array}
\]

Tableau (39) shows two candidates which satisfy both of the NP-movement constraints. Since both wh phrases are internal arguments, they are equally superior in the syntactic tree. This means that NP-moving either one of them to SpecIP satisfies SUBJ_SUP. Additionally, since both are human, SUBJ_HUM is also satisfied. Whichever one is NP-moved to SpecIP then satisfies SORTSUP. The tie on all constraints predicts that two optimal, and therefore attested, forms result.
Tableau (40) shows the same structure as (39), but without any human wh phrases. Here, too, there is more than one optimal form due to a tie in each constraint. Note that "tie" does not just mean "both candidates satisfy"—the two candidates actually both fail SUBJ HUM. They nonetheless tie (i.e., fare equally) on both constraints, forcing there to be more than one attested form.

In our last tableau there is the same structure again as in (39) and (40), but with only one human wh phrase.

To summarize this section, while each of the constraints in (31) through (33) is required in order to yield the correct output, it is nonetheless impossible to rank these three constraints relative to each other. This is because the classic "kitty-corner" distribution of check marks and stars, shown in (42), is absent. We therefore leave the three constraints unranked.

5.3 Differentiating Bulgarian from English from Chinese

How then do we differentiate Rudin's [+ Multiply Filled SpecCP] languages—Bulgarian and Romanian—from those which front only one wh phrase, like English? Moreover, how can these two
language groups be differentiated formally from languages like Chinese, which front no wh phrases? We adopt the model in Grimshaw (1993b), which makes use of the following constraints:

(43) \textbf{STAY} [Grimshaw (1993b:1)]

Star trace (= the Economy of Derivation).

(44) \textbf{OPSPEC} [Grimshaw (1993a:1, 1993b:1)]

All operators must be in some Spec position.

Grimshaw also suggests that "either unmoved wh phrases are not Operators, as assumed in work on the 'Wh Criterion' ... or we must revise OPSPEC to, say, TOPSPEC" [Grimshaw (1993b:15)]. We follow the former lemma of her suggestion, defining only the sorting key as "operator"; all other (dependent) wh phrases are nonetheless required to be in the same specifier position as the sorting key at overt syntax. We define the constraint as follows:

(45) \textbf{DEPSPEC} Any wh phrase dependent on the sorting key must adjoin to the sorting key in overt syntax.

With the constraints in (43) through (45) a crude typology is possible: STAY is ranked so highly in Chinese as not to allow even one wh phrase to front (cf. Huang 1982 for the details). We therefore posit the ranking in (49):

(49) \textbf{Chinese:} \quad \text{STAY} \succ [\text{OPSPEC}, \text{DEPSPEC}]

The English-Bulgarian difference is that English fronts just one wh phrase, the sorting key, while Bulgarian fronts them all:

(50) \textbf{English:} \quad \text{OPSPEC} \succ \text{STAY} \succ \text{DEPSPEC}

(51) \textbf{Bulgarian:} \quad \{ \text{OPSPEC} ; \text{DEPSPEC} \} \succ \text{STAY}

A final note on the other Slavic languages, which Rudin (1989) posits as having one wh word in SpecCP, with the rest perhaps adjoined to IP. We would essentially classify these languages with English, which fronts one wh phrase. We further
speculate, following the spirit of Yokoyama’s (1986) treatment of Russian, that the remaining wh phrases, being inherently “low-content” information, are fronted as part of the discourse-influenced constituent order, often referred to as “theme-rheme” structure. That is, all the Slavic languages except Bulgarian only have single syntactic wh-fronting. This also explains why such languages cannot front more than one wh phrase to a higher clause.\(^\text{17}\)


Notes:

1. The following abbreviations are used in this paper: ACC[usative]; ADV[verb]; Agro(P): object-agreement (phrase), a functional category; C(P): complementizer (phrase); CL[tic]; DAT[ive]; DO: direct object; F[eminine]; I[inflection] (a functional category); IP: inflection phrase; IO: indirect object; M[asculine]; N[euter]; NP: noun phrase; NPI[ction]; NOM[ative]; P(P): preposition (phrase); PF: phonetic form; PL[ural]; Pr(P): predicate (phrase); SG: singular; Specifier; f[rame]; V(P): verb (phrase); wh: (non-yes/no) interrogative; X\(^{(a)}\): position(s); XP: any maximal projection; 1, 2, 3: first-, second-, third-person.

2. Rudin (1988b:461-2) also uses clitic placement to contrast Bulgarian from the rest of Slavic, showing the contrast in (i) and (ii):

(i) **Bulgarian**

*Koj kakvo na kogo e dal?*  
who what to whom CL gave  
NOM ACC DAT 3.SG M.SG

(ii) **Serbo-Croatian**

*Ko je što kome dao?*  
who CL what to-whom gave  
NOM 3.SG ACC DAT M.SG

‘Who gave what to whom?’ (for both)  
[= ex. (29, 31), Rudin (1988b:462), resp.]

Only these orders are allowed (except for switching the non-initial wh words in (ii)). Rudin (1988b) points out that these data constitute a valid argument only against multiply filled SpecCP in Serbo-Croatian, not for multiply filled SpecCP in Bulgarian. Bulgarian e must always procliticize to the tensed verb, so its position is irrelevant for this test. Cf. also Avgustinova (1994) and Dimitrova-Vulchanova (1992) regarding the verb-proclitic status of Bulgarian auxiliaries and pronouns.

One Bulgarian clitic that does not require procliticization to the verb, *li* (y/n question), must nonetheless follow all wh phrases. [This is a correction of ex. (25) in Rudin (1994), which incorrectly reports ‘*Koj li kako u udaril?* Who on earth
killed whom?"; the use of *li in wh questions renders an emphatic, 'on earth' or 'the hell' meaning, and is difficult to get in multiple-wh questions. But if this question were uttered, it would be Koj kogo *li e udari?"

Another possible distinction between Bulgarian and the rest of Slavic is the prosodic phrasing of multiply fronted wh phrases. Cichoński (1985:58, citing C. Dogil p.c.) reports that in Polish when more than two wh phrases appear clause-initially there is an intonational boundary between the first wh phrase on the one hand and all the subsequent wh phrases. Our preliminary (non-instrumental) tests suggest that there is no such intonational boundary in Bulgarian wh clusters.

"No rule can involve X, Y in the structure ...X...{...Z...- WYZ...}... where the rule applies ambiguously to Z and Y and Z is superior to Y." [Chomsky (1975:246)]

In §5 we propose constraints that (a) require SpecIP to be occupied by the highest constituent (namely: the external argument) and (b) require SpecIP to be occupied by a human constituent. Whereas these constraints are not crucially ranked in Bulgarian, (12) and (13) would rank SUBISUP > SUBJHUM for English.

Unless otherwise noted, all examples were elicited by us.

Specifically, judgments depend on whether wh phrases are S- or VP-adverbials.

In this paper we do not consider multiple-wh questions with only adverbial wh phrases, aside from the preliminary observation that such questions tend to require the conjunction i 'and' between the two clause-initial wh phrases. [Cf. also n. 17.]

Dimitrova-Vulchanova also reports the order Na kogo kaj kakvo kaza? [her ex. (80c)], which was not in the earlier draft of her paper. Our informants judge this order to be "awkward" at best, not used in neutral contexts. She does add that "the only preferred position within the sequence is that of the subject wh-constituent" [p. 45].

Additionally, we offer yet another phenomenon: In the construct One can't not go to work, you know!, it's impossible to replace the contraction can't with the separate words can not: *One can not not go to work, you know!, primarily because of the consecutive *not not homophones.

The careful reader may have noticed that we have not presented any examples of a non-human wh external argument with any of the other wh phrases being non-human. Two such examples come to mind: First, there is the question with NOM and ACC non-human arguments, both of them kakvo due to the ambi-case status of this wh word. It is impossible to determine which kakvo is first in the wh cluster. (If both of these wh phrases appear clause-initially, then there is the additional factor of consecutive homophony: Kakvo kakvo e udari? 'What hit what'? in fact such an example was presented during questions following our talk at FASL-3). Second, there is the wh question with kakvo as the external argument as well as a wh adverbial. We have elicited the following:

(i) Kakvo kade te udari? (ii) Kade kakvo te udari?

what where you hit where what you hit

NOM ADV ACC 3.SG ADV NOM ACC 3.SG

Our informants tend to favor (ii). We have no explanation for this.

We assume, non-critically, that the NOM argument in (17) is generated as the external argument, while that of (18) is generated VP- Internally.

Kano & Takami (1993) have modified Kuno's original (1982) wording of this hypothesis using the word "leftmost" instead of "fronted", in order to account for wh adverbials which they posit as being base-leftmost-generated. For our purposes their revision likewise serves a purpose, distinguishing the very first wh phrase from any subsequent ones (in a multiply fronted wh cluster). The wording in (29) also serves to unify Bulgarian with single-wh-fronting languages like English; in each the first or
only "leftmost" wh phrase is the sorting key. We leave aside until §5.3 why languages like Chinese (and Japanese) do not move any wh phrases overtly.

Unfronted wh phrases can also be dependent on higher-clause wh sorting keys or be “discourse-linked” (see Williams 1994:191-195 and Pesetsky 1987). We limit this paper to context-free wh questions not embedded in other wh clauses.

Our present proposal is not crucially inconsistent, in most respects, with the earlier assumptions in Rudin (1988b; 1989), that all wh phrases are fronted to SpecCP. Crucially, our present model requires (based largely on the arguments above in §1) that all wh phrases be in the Spec of the topmost projection (i.e., SpecIP or SpecCP) in overt syntax (or, in the case of wh-extraction, as in (3), the same destination-specifier position). Both Grimshaw’s and Rudin’s proposals agree on one environment: If a non-wh element occupies SpecIP, then a CP projection is required. In such a case tensed V (with its clitics) moves to C. Izvorski (1993) proposes yet a third configuration: all interrogative wh phrases are fronted to the Spec of a split-l projection [F[ocu]s][P]; that is, no interrogative wh phrases are fronted to SpecCP (but relative-clause wh phrases are).

One argument in favor of Grimshaw’s model is the following: Note that the summary at the beginning of §5 is not entirely straightforward (“If there is no external argument of any kind ... and if there is only one human wh phrase, as in (15) and (18)–(19), then that human wh phrase must appear first in the wh cluster.”). Actually, the “only ... human wh phrase” in (15) or (18) is the only human argument of any kind. In (19), however, there is also another human argument, the INDEF clitic pronoun ti. Recall as well that SUBJHUM does not discriminate as to wh-hood. If ti is NP-moving to SpecIP instead of koj ‘who’ in (19), then SUBJHUM (and SURFIC) would be satisfied and then both wh phrases would be movable to SpecCP, with either one as sorting key satisfying SORTSUP. This appears to be true. In the overall assessment, however, NP-moving koj to SpecIP satisfies more constraints than if ti were so moved, as shown in the following structures:

(i) Spec[koj; začu] [NPi ...whfi ...ti] [li_j i_1] [CP]
(ii) Spec[koj; začu] [NPi ...whfi ...ti] [li_j i_1] [CP]
(iii) Spec[koj; začu] [NPi ...whfi ...ti] [li_j i_1] [CP]
(iv) Spec[koj; začu] [NPi ...whfi ...ti] [li_j i_1] [CP]

All four structures satisfy each of SURFIC, SUBJHUM, and SORTSUP. SUBFIC is satisfied equally, since all constituents movable to SpecIP are equidistant to it, all within VP; SUBJHUM is satisfied in each because either koj ‘who’ or ti ‘you’ is NP-moving to SpecIP; finally, SORTSUP is satisfied in each because a) if koj is selected to be in SpecIP (prior to any wh-movement), as in (i) and (ii), then koj is also leftmost in overt syntax; or b) ti is in SpecIP, as in (iii) and (iv), and either wh phrase is equidistant to SpecCP, satisfying SORTSUP equally.

How then is (iv)—the unattested order—ruled out? Stay, in (43), selects (i) over any of (ii) through (iv). While it is not clear how stay is assessed, it is obvious that example (i), with only two movements (one wh- and one NP-movement), violates stay less than any of (ii), (iii) or (iv) (which each undergo one extra wh-movement). Thus, the unattested order in (iv) is ruled out. In addition, the fact that (i) is more optimal than either (ii) or (iii) is also an argument in favor of Grimshaw’s (1993a) build-a-CP-only-when-necessary model (cf. fn. 16 and 17).
While we do not show the tableaux for (20) and (21) here, their distribution of stars and checks is the same. We assume that the inaudible external argument pronoun in (20) and (21) behave like Ivan in (16) and (37) (in the relevant respects).

In the tableaux, we use whichever register (colloquial or literary) that conveniently eschews any obfuscation caused by consecutive homophony. Note that (38a-b) correspond to (16a-b) respectively, and that (16b) is bad, due to STARHOM (cf. §4). This is controlled for in their colloquial counterparts in (27).

Recall from (16) that Ivan is clause-final on the surface. We rely on V-to-C movement to yield the post-verbal position of Ivan. V-to-C occurs, however, only when a CP is formed, which, under Grimshaw's and our account, (cf. nn. 14 and 17) only happens when SpecP is non-wh. This supports Grimshaw (1993a) even more.

Since circulating presenting this paper and circulating it in 1994 we have learned of two other treatments of Slavic multiple-wh phenomena: Golden (1995) reports multiple wh-extraction in Slovene; Przebłiórkowski (1994) also finds that multiple adjunct wh phrases in Polish must be conjoined using it ‘and’ and reports Polish data with the complementizer te ‘that’ which suggest (to us) a SpecP analysis for the subject wh phrase, along the lines of Grimshaw (1993a; 1993b); cf. nn. 14 and 16. We have also learned of the following two Optimality-theoretic approaches to the three-way wh-fronting typology in (49) through (51): Ackema & Neeleman (1995) and Legendre et al (1995). All four works have apparently arrived at their conclusions independently. We refrain from entertaining these other works' arguments here. For posterity, however, we list these works below.

References:


Rudin, Catherine (1985) "Multiple wh-movement and the Superiority Condition." Ms., Dept. of English, Miami Univ., Ohio, read at LSA.

Rudin, Catherine (1986) *Aspects of Bulgarian syntax: Complementizers and wh constructions.* Columbus: Slavica.


Rudin, Catherine (1994) "On focus position and focus marking in Bulgarian questions." In *Papers from the Fourth annual meeting of the Formal Linguistics Society of Midamerica.* A. Davison et al (eds.). Iowa City: Dept. of Linguistics, Univ. of Iowa. 252-265.

Slabakova, Roumyana (1994) "Bulgarian psyche verbs." This volume.


Vakareliyska, Cynthia (1994a) "Na-drop in Bulgarian." *Journal of Slavic linguistics* 2:1, 121-150.

Vakareliyska, Cynthia (1994b) "Evidence from aphasia for a subject topic slot in Bulgarian." This volume.


billings@mailer.fsu.edu crudin@wscgate.wsc.edu

Modern Languages and Linguistics
Florida State University
Tallahassee, FL 32306-1020

Humanities Division
Wayne State College
Wayne, NE 68787