

Topicality and Superiority in Bulgarian *wh*-questions

T. Florian Jaeger, Stanford University

1 In a nutshell*

In this paper, I present a topicality-based analysis of object clitic doubling in *wh*-questions in colloquial Bulgarian. The analysis draws intimately on earlier research on two phenomena well-known in the literature on Slavic syntax, Superiority (Chomsky 1973) in Bulgarian multiple *wh*-questions (e.g. Billings & Rudin 1996, 1998; Bošković 1993, 1998a,b; Citko & Grohmann 2001; Grewendorf 2001; Lambova 2003; Pavlov 2000; Richards 1997; Rudin 1985, 1988a,b, 1989; see Blaszcak & Fischer 2001 for a recent overview of the literature) and clitic doubling of topic-fronted objects (e.g. Alexandrova 1997; Avgustinova & Andreeva 1999, Guentchéva 1994; Jaeger 2002; Jaeger & Gerassimova 2002; Leafgren 1997, 1998; Rudin 1997; Schick & Beukema 2001). (1) illustrates the effect of Superiority. The subject *wh*-word *koj* ‘who’ has to precede the direct object *wh*-word *kogo* ‘whom’. (2) is an example of clitic doubling of a fronted phrase (henceforth CD).¹ Note that CD is possible out of embedded clauses as well as within a clause (no example of the latter is given here).²

* I am very grateful to T.H. King, I. Sag, L. Billings, and K. K. Grohmann for their critical and detailed feedback on theoretical issues. I am also extremely grateful to V. Gerassimova, M. Dimitrova-Vulchanova, and M. Tasseva-Kurktchieva for discussion of critical empirical issues. Finally, I want to thank the anonymous FASL-12 reviewers, the audience of FASL-12 - especially O. Arnaudova, and M. Lambova – for very useful and supportive comments, as well as the audience of an earlier presentation at Stanford, and my informants. None of the above-mentioned people necessarily agrees with the views expressed here and, as always, all remaining mistakes are mine.

¹ Here, I am not interested in the difference between clitic doubling of ‘topicalized’ phrases and clitic left dislocation (but see Arnaudova 2002; Jaeger 2002). All that matters to current approach is that the same mechanism involving CD available in declarative clauses is also available in *wh*-questions, as I will show below.

² I use the following mark-up conventions in examples: fronted constituents such as *wh*-words and topics are underlined; object clitics and the fronted constituents co-referential to them are doubly underlined; object clitics are *italicized*. When necessary, I use subscripts in the literal translation to annotate person (1, 2, 3), number (SG, PL), gender (MASC, FEM, NEUT) and/or definiteness (+/-DEF).

Superiority in Bulgarian *wh*-questions: subject > object

- (1) a. Koj kogo kak e celunal?
who whom how is kissed
b. *Kogo koj kak e celunal?
Who kissed whom how?

Clitic doubling of a topic-fronted object [Jaeger 2002:4]

- (2) Todor e jasno, [če Ivan *(go) e vidjal].
Todor is clear that Ivan him is seen
Todor, it is clear that Ivan has seen (him).

Although both phenomena have received an enormous amount of attention in the literature on Bulgarian as well as in cross-linguistic studies (see above for Bulgarian; for Superiority and CD in other languages, see Jaeger 2003:183), no account has yet been presented of CD in *wh*-questions, as in (3), - that is the only construction which falls in the domain of both Superiority *and* CD.

CD in *wh*-questions [Dimitrova-V. & Hellan 1998:xxii]

- (3) Na kogo kakvo mu dadoxa?
to whom what IOC gave_{3,PL}
To whom did they give what?

While CD has been connected to D-linking before, I focus on new data containing CD of *bare wh*-phrases. Based on the new data, I argue that topicality is a determining factor for *wh*-phrase ordering in Bulgarian multiple *wh*-questions.³ By tying *wh*-phrase ordering to topicality rather than directly to CD, the data presented below bears on *wh*-phrase ordering in general (i.e. in questions with or without CD). This in turn raises the question to which extent topicality accounts for the so-called Superiority effects.

The remainder of the paper is organized as follows. Section 2 gives a brief overview of the role of Superiority in Bulgarian *wh*-questions and summarizes two prominent competing proposals. In

³ For all practical purposes in this paper, readers more familiar with the syntactic literature than the literature on information structure can substitute '(overt and covert) D-linking' for 'topicality'. Note, however that the two concepts are not identical. I will return to this issue in section 5.

section 3, I introduce new data showing CD in multiple *wh*-questions. Section 4 provides a brief introduction to the literature on CD in non-question contexts and links it to topicality of the object. This is the point of departure for section 5, where I bring together the previously unconnected research discussed in sections 2 and 4 and develop an account for CD in Bulgarian *wh*-questions. Finally, section 6 summarizes the argument made in this paper and discusses its implications and possible avenues for further research.

2 Multiple *wh*-questions and Superiority

In this section, I summarize the two accounts of multiple *wh*-questions in Bulgarian, that are most relevant to the remainder of the paper, one by Bošković (1998a,b) and one by Billings & Rudin (1996, 1998).

In Bulgarian non-echo questions, all *wh*-phrases are obligatorily fronted (see (1) above) and can only be preceded by topic-fronted phrases, an example of which is given below:

Topics precede *wh*-words [Rudin 1985:92]

- (4) Ivan na kogo dade knigite?
Ivan to whom gave books_{+DEF}
Ivan, to whom did (he) give the books?

Bulgarian multiple *wh*-questions have been the subject of an ongoing debate under the label Superiority (see references in section 1). That is, the goal of the research on multiple *wh*-questions was to understand the determining factor(s) for the ordering of fronted *wh*-phrases. With the emergence of the Minimalist Program, the debate has received even more attention. In recent years, several proposals have been made to account for the apparent Superiority effects in Bulgarian (e.g. Billings & Rudin 1998; Bošković 1998a,b; Grewendorf 2001; Richards 1997).

Revising an earlier claim by Rudin (1988a,b, 1989), Bošković (1993, 1998a,b) shows that Bulgarian *wh*-questions show selective

Superiority effects. That is, the first *wh*-phrase is subject to Superiority, as in (5) and (6), and all following *wh*-phrases can order freely after the initial one, as shown in (7).

Subject > direct object

- (5) a. Koj kogo e tselunal?
who whom is kissed
b. *Kogo koj e tselunal?
Who kissed whom?

Direct Object > indirect object [Lambova 2003]

- (6) a. Kakvo na kogo e kazal?
what to whom is said
b. *Na kogo kakvo e kazal?
What did he say to whom?

Subject > direct object <> adverb [based on Bošković 1993]

- (7) a. Koj kogo kak e tselunal?
who whom how is kissed
b. Koj kak kogo e tselunal?
Who kissed whom how?

Bošković's hypothesis (1998a,b) accounts for these ordering facts by means of economy and thereby reduces Superiority to a merely "descriptive generalization" (Bošković 1998a:60).

Hypothesis 1: Bošković (1998b:7)

"Superiority effects follow from the requirement that the strong +*wh* feature of C be checked in the most economical way."

A number of exceptions to this generalization have been reported in the literature. For example, Grewendorf (2001:97) further constrains Superiority to subjects, and shows that Superiority is not obeyed by non-subject initial *wh*-words. This is also supported by some examples in Billings & Rudin (1996):

Direct object <> indirect object [Grewendorf 2001:97]

- (8) a. Kakvo na kogo e dal Ivan?
what to whom is given Ivan
b. Na kogo kakvo e dal Ivan?
What has Ivan given to whom?

Direct object <> adverb [Billings & Rudin 1996:42]

- (9) a. Kakvo koga e kupil Ivan?
what when is bought Ivan
b. Koga kakvo e kupil Ivan?
What did Ivan buy when?

Billings & Rudin (1998) provide the following examples against Bošković's (1998b) model:⁴

Subject [-hum] <> direct object [+hum] [Billings & Rudin 1998:5]

- (10) a. Kakvo kogo e udarilo?
what whom is hit
b. Kogo kakvo e udarilo?
What hit whom?

Int. subject [-hum] <> ind. object [+hum] [Billings & Rudin 1998:6]

- (11) a. Koj na kogo mu xaresva?
who to whom ioc pleases
b. Na kogo koj mu xaresva?
Who likes whom?

According to Billings & Rudin (1996, 1998), (10) is due to animacy. That is, a [+hum] accusative case marked *wh*-phrase and an inanimate external subject *wh*-phrase can occur in both possible orders. (11) contains an “impersonal psych predicate”, a “di-unaccusative” (Billings & Rudin 1998:6) with two internal arguments. The dative marked *wh*-phrase and the internal nominative marked *wh*-phrase can order freely because both are [+hum]. Both cases serve to license violations of Superiority. Considering many other configurations, Billings & Rudin (1996:46) conclude:

Hypothesis 2: Billings & Rudin (1996:46)

External subject *wh*-phrases must be the first *wh*-phrase. If there is no external argument, then animate *wh*-phrases must precede inanimate *wh*-phrases. All other *wh*-phrases order freely.

⁴ ioc refers to the indirect object clitic; DOC refers to the direct object clitic. A brief introduction to object clitics is given in section 4.

Next, I will introduce new data, similar to (11) involving CD. Following that I very briefly summarize the research on Bulgarian CD outside of questions and link the two phenomena to each other.

3 CD in *wh*-questions

In this section, I introduce new data, thereby extending the observations made by Billings & Rudin (1998) to a more general class of examples that involve CD in *wh*-questions. If not stated otherwise, the data were collected by means of surveys and interviews with several native speakers (some linguistically naïve, others linguistically trained). Since CD is restricted to colloquial speech, informants were given instructions accordingly (i.e. ‘Could you say that while talking to your friends?’).⁵

Billings & Rudin (1998:6) restrict examples like (11) above to di-unaccusatives. However, as shown in (3) above, in which the indirect object of ‘give’ (clearly not an unaccusative) is CDed, CD of *wh*-words is possible in other contexts, too (Dimitrova-Vulchanova & Hellan 1998:xxii).

Jaeger (2002:31) mentions an example of CD in a question, here given as (12), where the CDed direct object *wh*-phrase precedes the subject *wh*-phrase.⁶

CD: direct object > subject [Jaeger 2002:31]

- (12) a. Kogo kakvo go iznenada?
whom what DOC surprised_{3.SG}
b. ?*Kakvo kogo go iznenada?
What surprised whom?

Note that, in contrast to the examples discussed in the previous section, the CDed *wh*-word in (12) *must* be fronted, whereas in

⁵ Note further that availability of CD differs among dialects of Bulgarian (as a rule of thumb, dialects spoken closer to the Macedonian border are more prone to CD). All informants I interviewed had CD in declaratives and questions.

⁶ It is not clear whether Billings & Rudin (1998) would have analyzed *kakvo* ‘what’ as an internal argument, but in this case further explanation would be necessary since *iznenadvam* ‘to surprise’ can also take [+hum] (agentive) subjects.

Billings & Rudin's (1998:6) example of a [-hum] subject (in this case an external argument), given above as (10), as well as in the CD example in (11), both orders of subject and object were possible.

Interestingly, CD of *wh*-words is generally not restricted to specific verb types in colloquial Bulgarian. (13) is another example with a [-hum] external argument and (14) shows that the CD of fronted *wh*-phrases is possible (and obligatory) for [+hum] external arguments as well.⁷

CD: direct object [+hum] > subject [+/-hum]

- (13) a. Kogo kakvo * (go) ubi?
whom what DOC killed
- b. Kakvo kogo (?go) ubi?
What killed whom?
- (14) a. Kogo koj * (go) narisuva/pozna/vidja?
whom who DOC painted/recognized/saw
- b. Koj kogo (?go) narisuva/pozna/vidja?
Who painted/recognized/saw whom?

To further illustrate the strength of the generalization that the CDED *wh*-phrase must be fronted, note that my informants (quite strongly) rejected (15) because it “didn’t make sense” to them (but see (16)).⁸

CD: semantic oddity due to enforced interpretation

- (15) #Kogo kakvo go običa?
whom what DOC loves
[Translation by informants] *What loves whom?*

⁷ Judgments differed slightly (across informants and across different sessions with the same informant) in how well-formed (b) was considered with the DOC. Judgments sometimes seemed to depend on the verb, although the verbs did not pattern together according to any known lexical class (e.g. aspect). It is also important to note that some speakers seem to lack any Superiority restrictions even without CD (Mila Tasseva-Kurktchieva, *p.c.*). Idiolects lacking Superiority in all cases were not considered here.

⁸ Note that the fronted phrase is interpreted as object if the more colloquial direct object *wh*-word *koj* (cf. Engl. ‘who’) is used. In this case, nothing should prevent the interpretation ‘Who loves what’ since *koj* can also be subject *wh*-word. The fact that the nonsensical reading ‘What loves whom’ is preferred even then (Veronica Gerassimova, *p.c.*) supports my point.

- (16) Kogo koj go običa?
 whom who DOC loves
Who loves whom?

Given (13) - (16) it seems that CD prefers or even *requires* the doubled *wh*-phrase to be fronted, thereby apparently violating Superiority even for the subject, and even if the external argument is [+hum], contrary to both Bošković's and Billing & Rudin's hypotheses. Furthermore the class of examples in (13) - (16) does not require psych-verbs or di-unaccusatives.

This raises the following questions. Do the CD-variants (with the fronted object *wh*-word) and the non CD-variant of a *wh*-question have the same meaning?⁹ Do they have they same distribution/uses? If not, what does the CD-variant mean?

Before addressing these questions (in section 5), I will set the ground for an analysis of CD in questions by providing a very brief introduction to CD in non-questions, and linking CD to topicality.

4 Clitic doubling and topicality in non-interrogatives

Like all other languages of the Balkan-Sprachbund, Bulgarian allows for CD by either the indirect object clitic (IOC) or the direct object clitic (DOC). The CDed phrase does not have to be fronted, e.g. (17), but for fronted topical objects CD is obligatory, as shown in (18).

Optional doubling of non-fronted objects [Jaeger 2002:3]

- (17) Decata (ja) običat Marija/neja.
 Children+DEF DOC_{3.SG.FEM} love_{3.PL} Maria/her_{3.SG.FEM.ACC}
The children love Maria/her.

Obligatory doubling of topic-fronted objects [Jaeger 2002:3]

- (18) Marija/Neja * (ja) običat decata.
 Maria/her_{3.SG.FEM.ACC} DOC_{3.SG.FEM} love_{3.PL} children_{DEF.PL}
Maria/her, the children love.

⁹ Up to this point, I have tacitly assumed this in the translations only for the sake of simplicity.

Many hypotheses about the function of CD in Bulgarian have been proposed (see Jaeger 2002 for an overview). Nevertheless, it seems fair to say that there is now considerable agreement that CD of fronted constituents marks *topicality* (for references, see section 1). More specifically, I have argued elsewhere (Jaeger 2002; Jaeger & Gerassimova 2002; see also Leafgren 1997) that this holds if topicality is understood in the sense of ‘aboutness’ (Reinhart 1982, Lambrecht 1994), as illustrated in the following example, where the sentence is intended to provide information about ‘the elephants’:

CD: Topic-marking [Jaeger 2002:16]

- (19) Slonovete_{TOP} * (gi) obučavat xorata.¹⁰
 elephants_{+DEF} DOC_{3.PL} train_{3.PL} people_{DEF}
The elephants, (the) people train (them).

Few have mentioned that CD is also possible in questions (but see Dimitrova-Vulchanova & Hellan 1998; Jaeger 2002), and nobody has yet discussed the meaning/usage of such examples.

5 Putting it together: Topicality in *wh*-questions

I propose that topicality is the driving factor for CD of *wh*-phrases. Thus CD of *wh*-phrases follows the same rules that are well-known for CD in declaratives. This hypothesis is formulated below:

Hypothesis 3: WH-Topic-fronting Hypothesis (WTH)

- (A) CD in Bulgarian *wh*-question indicates that the CDed *wh*-phrase is the topic of the question.
 (B) Thus fronting of CDed *wh*-phrases is due to the same feature (topicality) that causes topic-fronting in non-question clauses.

Ad (A), what does it mean to be the *topic of a question*? *Wh*-fronting is commonly associated with *focus* rather than topicality

¹⁰ (19) is grammatical without the DOC if *slonovete* ‘the elephants’ is realized with emphatic stress, in which case it is interpreted as exhaustive focus.

(see e.g. Bošković 1998a,b). This view is usually based on either or both of the following two observations: (a) *wh*-words request *new information*, and (b) like focused phrases, *wh*-words can be treated as *variables in the proposition* expressed by the sentence they occur in.

Here, I adopt a more than one-dimensional approach to information structure, in which topicality is defined independently of focus. Thus a semantic object may be focused and topical at the same time (cf. Dik 1989:Ch. 13; Jaeger & Oshima 2002:13-14; Lambrecht 1994).¹¹ Ergo, all fronted *wh*-phrases are foci and sometimes a *wh*-phrase can also be the topic of the question (cf. Gundel 1988:210; Leafgren 1997:127; Steedman 2000:659).

In the approach taken here, the notion ‘topic of a question’ is closely related to D-linking (Pesetsky 1987) and I will come back to the relation between these two concepts shortly. First, I present an example that is representative for contexts licensing CD of *wh*-phrases. I chose verbs with a prototypical agent and patient, and asked my informants in what context the verbs could occur in a CD *wh*-question. Consider the context in (20) which licenses the question in (21):

Context facilitating topicality of the direct object in (21)

(20) *Some of the most popular painters in town recently made portraits of a couple of my friends. I know that each of my friends wanted to be painted by a particular artist, but I don't know by whom. So the question is :*

(21) *Kogo koj go e narisual?*
whom who DOC has painted
Who has been painted by whom?

The question in (21) was judged to be felicitous only in a context in which the set of friends has been mentioned before *and* is salient in the discourse. This suggests that ultimately it will not be enough to only require CDed *wh*-phrases to be D-linked (in the sense that

¹¹ Within the literature on information and discourse structure, the necessity to account for ‘new/focused topics’, as in e.g. the English left-dislocation (cf. Keenan-Ochs & Schieffelin 1976), has been recognized for some time.

the set of referents of the *wh*-phrases are ‘given’ in the discourse; cf. Pesetsky 1987:107f.).¹² Note, however, that CDed *wh*-phrases seem to be a subtype of D-linked items both syntactically (whereas D-linking frees *wh*-phrases from Superiority, CDed *wh*-phrases have to be fronted; see section 3 and below) and functionally (CDed *wh*-phrases have to be *very* salient in the discourse).

This leads me to assume the stronger hypothesis that the WHTH should be interpreted in terms of an aboutness-based approach (e.g. Jacobs 2001; Lambrecht 1994; Reinhart 1982), in which the topic of a question is ‘what the sentence primarily requests information about’ (cf. Gundel 1988:210; Leafgren 1997:127; see also Kuno’s *key sorting hypothesis* in section 6.1). In such an approach, topicality implies (covert or overt) D-linking because, roughly speaking, whatever ‘Xs’ a piece of discourse is ‘about’, those ‘Xs’ are ‘given’ in the discourse and therefore D-linked. However, not everything that is D-linked is also a topic (in the aboutness-based sense). The idea that the CDed *wh*-phrases are aboutness-based topics (rather than merely D-linked) captures the intuition of my informants that (20) and (21) are *about* the ‘friends’. I take this to support clause (A) of the WHTH.¹³

Clause (A) makes strong predictions about the distribution of CDed *wh*-phrases, which follow from the distribution of topic-fronted non-*wh*-phrases in questions and non-question clauses (cf. section 2 and 4). This is summarized in clause (B) of the WHTH, a corollary of (A). Next, I discuss some of the consequences of (B).

First, CDed *wh*-words should be able to precede D-linked *wh*-phrases, which have been shown not to be subject to Superiority (Rudin 1988a:476; Richards 1997:42-3; cf. Blaszczyk 2002:23-24), *even if the D-linked phrase is the subject*. This prediction is borne out, as shown in (22) where the direct object *wh*-phrase *kogo*

¹² See Comorovski (1996:2) for a definition of D-linking that incorporates saliency.

¹³ Here I do not have the necessary space to motivate an aboutness-based approach in more detail but I hope to have shown that D-linking (as defined in Pesetsky 1987) and CD of topical *wh*-phrases differ in their distribution. For cross-linguistic support for topicalization of *wh*-phrases, see e.g. Grohmann (1998) and Scott (2003).

precedes the D-linked subject *wh*-phrase *koja žena*. Note that this is even the case when the colloquial direct object *wh*-word *koj* is used, in which case nothing in the *wh*-phrases themselves marks *koj* to be the direct object (both *koj* and *koja žena* can be either nominative or accusative).

CD: direct object > D-linked subject

- (22) Kogo/koj koja žena * (go) običa?
 Whom/whom which_{FEM} woman DOC loves?
Who does which woman love?

Second, a similar type of evidence for the WTH comes from sentences with two inherently D-linked *wh*-phrases. In (23), the absence or presence of the DOC determines whether the first or the second *wh*-phrase is interpreted as object (inherently D-linked *wh*-phrases are not case marked). Interestingly, the same two types of readings are available even for (22) with the bare *wh*-phrase *koj*.

CD: determining the interpretation of D-linked object & subject

- (23) Koj máž koja žena (go) običa?
 Which_{MASC} man which_{FEM} woman DOC loves
 Without DOC: *Which man loves which woman?*
 With DOC: *Which woman loves which man?*

Third, assuming a topic-sensitive ordering constraint on the left periphery (cf. Jaeger 2003:294f.) more prototypical topics, such as topic-fronted lexical NPs, should precede CDed *wh*-words. This prediction is borne out. Whereas topic-fronted non-*wh*-phrases can precede topic-fronted *wh*-words (Lambova 2003:130f.; Rudin 1985:92; also cf. (4) above), the opposite order is not possible:

Fronted topics have to precede the *wh*-cluster

- (24) *Kogo, na Marija, koj (i) (go) e pokazal?
 whom to Maria who IOC DOC e shown
 Intended: *As for Maria, who showed her whom?*

Fourth, if CD really marks topicality and topicality is independent of focus-fronting, then CD of *wh*-words should not depend on the presence of multiple *wh*-phrases (although it is expected that CD is

much more common with multiple *wh*-questions, in which there is a need to identify the topic of a question – after all it is often the speaker’s choice to overtly mark a topic for the sake of clarity; cf. Leafgren 1997, a.o.). Like the first two predictions, this one is borne out as well. Example (25) was found in Aleksova’s corpus of colloquial Bulgarian (*k’o* is a colloquial form of *kakvo* ‘what’).¹⁴

CD of single *wh*-phrase [from Aleksova’s corpus]

[*The speakers are discussing different options for what they could knit: a jacket, a pullover, etc.*]

- (25) Kaži majko k’o da ti go oplita?
 tell mother what SBJfor-you DOC knit
Say, mother, what (shall I) knit for you?

Fifth, given that the echo *wh*-phrase of an echo-question is arguably the topic of that question (‘what the question primarily requests information about’), the fact that the echo *wh*-phrase in (27) is CDed can be taken to support the WHTH:

- (26) Koj e narisual Ivan?
Who has painted Ivan?
 (27) Kogo koj go e narisual?
 whom who DOC has painted
Who has been painted by WHOM?

I have given further support for the WHTH in Jaeger (2003:190f.). I argued that some elements (e.g. emphatic particles and certain adverbs) known to appear between topic-fronted NPs and the *wh*-cluster (Lambova 2003) also appear between CDed *wh*-phrases and the rest of the *wh*-cluster. If CDed *wh*-phrases are topic-fronted, this phenomenon is straight-forwardly accounted for.

To sum up, the predictions made by the WHTH about the distribution of CDed *wh*-phrases are borne out and the WHTH is supported by native speaker intuitions.

¹⁴ Krasimira Aleksova’s corpus consists of transcribed conversations in family contexts (approx. 138,000 text words; <http://www.hf.uio.no/east/bulg/mat/Aleksova/>). I am grateful to Olga Arnaudova for making me aware of the preceding context.

6 Summary, Implications and a new perspective

I have presented new data from colloquial Bulgarian relating the research on (multiple) *wh*-questions to the research on clitic doubling (CD). I have argued that CD marks topicality (see the WHTH formulated in section 5) – within as much as outside of *wh*-questions. I have done this on the basis of three types of evidence. First, the interpretation and translation by native speakers was taken to ascertain if and how the meaning of CD questions is different from non CD questions. Second, the distribution of CD questions, or more precisely the fact that they are licensed by a specific type of context, shows that only topical object *wh*-phrases can be CDed. Finally, I have shown that the predictions made by the WHTH are borne out.

Next, I will discuss several possible implications of the analysis proposed here (section 6.1), sketch how a formal account might look (section 6.2), and mention two important open questions (section 6.3).

6.1 *Implications & Outlook*

Strikingly, CDed *wh*-phrases *must* be fronted before non-CDed *wh*-phrases - just as topic-fronted elements in non-interrogatives must precede the focus-fronted elements. While I have shown that CDed objects are topical in both interrogatives and non-interrogatives, and that topical objects must be fronted, I have not yet said anything about topical *non*-object *wh*-phrases. However, the parallelism between CD in interrogatives and non-interrogatives, and the well-known fact that in Bulgarian declaratives at least some non-object phrases (i.e. subjects, non-object arguments, and adjuncts) can be topic-fronted without CD are compelling (there are no clitics for non-object arguments).

This raises the question to which *extent* topicality directly influences *wh*-ordering (rather than only indirectly through CD). Consider the following hypothesis which declares topicality rather

than CD to be the cause for ordering constraints on fronted *wh*-phrases:

Hypothesis 4: Topics-First! Hypothesis

Like order constraints on the left-periphery of non-interrogatives, *wh*-phrase ordering (including so-called Superiority effects) in Bulgarian *wh*-question is (partly) determined by topicality.

There is not enough space here to discuss this issue in appropriate detail – I discuss, however, very briefly (a) why I think it is worth investigating the validity of the Topics-First! Hypothesis, and (b) some points that support it.

Ad (a), reducing a linguistic phenomenon (in this case *wh*-phrase ordering) to another, more general and independently motivated principle, is preferable to stipulating a specific constraint such as Superiority. This leads me to (b). Firstly, topics precede foci in non-interrogatives, and CDed *wh*-phrases are topical. Second, cross-linguistically subjects have been observed to be most frequently topics (cf. Givón 1976, Lambrecht 1994:131f). This would be in line with both native speaker intuitions regarding multiple *wh*-questions with initial subjects (i.e. they seem to primarily request information about the subject), and the fact that, in most multiple *wh*-questions with a subject *wh*-phrase, the subject is the initial *wh*-phrase. Cross-linguistic support for topic-fronting in *wh*-questions comes from colloquial Russian (Scott 2003), German (Grohmann 1998, and upcoming work), and from Romanian (Comorovski 1996). Finally, the Topics-First! Hypothesis captures the spirit of Kuno & Takami's (1993:112) *sorting key hypothesis* and ties it to information structure, by claiming that the *key of a question* is the topic of that question.

Hypothesis 5: Kuno & Takami's (1993:112) *Sorting Key Hypothesis*

“In a multiple *wh* question, the leftmost *wh*-word represents the key for sorting relevant pieces of information in the answer.”

6.2 *A short note on a formal analysis*

Formally, the data introduced in section 3 can be captured by assuming that topics and foci are fronted to one and the same left-periphery field that is the locus of discourse functions of different kinds (cf. ΔP in Lambova 2003:8; see also the multiple-filler construction in Jaeger 2003:196). The correct order of fronted constituents (i.e. topics before foci) is achieved by positing a linear order constraint sensitive to topicality on this left-periphery projection. Such an approach is spelled out in detail in the HPSG analysis presented in Jaeger (2003) where a linear precedence constraint working on the left-periphery orders fronted constituents. The commonalities between CD in *wh*-interrogatives and non-interrogatives are captured via a type-hierarchy. Lambova (2003a:10), dealing with data related to the cases presented here, suggests that the order of fronted constituents may be due to prosodic constraints on the linear order between the intonational contours on topics and foci. If this turns out to be true, the order constraint would not have to be stipulated for information structure but would rather be the (indirect) effect of (more general) prosodic constraints and the information structure to prosody mapping.

6.3 *Open questions & future research*

If the Topic-First! Hypothesis turns out to be true, it immediately raises the question as to whether topicality is the *only* ordering factor on the left periphery or whether it is *one among several* factors (e.g. animacy, agentivity, and/or subjecthood). More specifically, a very relevant question would be whether there are *any* Superiority effects (as defined in Chomsky 1973; or Bošković 1998a,b) that cannot be reduced to other (non-syntactic) factors.

One directly related empirical question is whether the second and third *wh*-phrases order freely in non-echo questions with at least three *wh*-phrases and one CDed (and therefore initial) object *wh*-phrase. I have not yet been able to determine the acceptability

of (28a) and (28b) with or without the indirect object clitic (IOC) *mu* to a satisfying degree:

Are both orders acceptable and, if yes, do they require the IOC?

- (28) a. Koja kniga koj na kogo (mu) ja dade?
which book who to-whom IOC DOC gave
- b. Koja kniga na kogo koj (mu) ja dade?
Intended: *Which book did who give to whom?*

A second empirical question is that apparently *some* speakers of Bulgarian do not at all allow CD of *kakvo* ‘what’, while it is marked, but acceptable for others (see also (25) above). The status of *kakvo* cannot be due to a general constraint on the fronting of inanimate *wh*-phrases, as suggested in Billings & Rudin (1996, 1998; cf. section 2 above), because CD of inanimate object *wh*-phrases is grammatical, as shown in (28). Interestingly, similar restrictions have been observed for other languages (e.g. Pesetsky 1987 for English who notes that ‘what’ is preferably *not* D-linked although it can be). This suggests that D-linking (as a syntactic notion) is a prerequisite for topic-fronting. The precise nature of the relation between D-linking and aboutness-based topicality is subject to further research.

Another avenue of research that I take to be crucially necessary investigates the semantics of multiple *wh*-interrogatives and especially CDed *wh*-questions. Kazenin (2000; cf. Blaszcak 2002:24) argues that *wh*-questions with (contrastive) single-pair interpretations obey Superiority but pair-list *wh*-questions do not show Superiority effects (see also Comorovski 1996:129ff. for the related distinction between *conjoined* and *matching* questions). It remains to be seen whether this distinction, if it turns out to be true, follows from more general constraints on the syntactic effects of semantic scope in Bulgarian or whether is specific to *wh*-fronting. If the former holds, Superiority could be a side effect of scope (cf. Bošković 2001 for a recent investigation of the relation between Superiority and the semantics of multiple *wh*-questions).

7 References

- Alexandrova, G. 1997. Pronominal Clitics as g(eneralized) f(amiliarity)-licensing AGR⁰. In *Proceedings of FASL: The Cornell Meeting 1995*, 1-31. Ann Arbor: Michigan Slavic Publications.
- Arnaudova, O. 2002. Clitic Left Dislocation and Argument Structure in Bulgarian. In *Proceedings of FASL-10*, University of Michigan.
- Avgustinova, T. 1997. *Word Order and Clitics in Bulgarian*. Saarbrücken Dissertations in Computational Linguistics and Language Technology, Volume 5. Saarbrücken: DFKI.
- Billings, L. & C. Rudin. 1998. Animacy and focus in Bulgarian wh questions. In *Proceedings of Comparative Slavic Morphosyntax*, Bloomington, Indiana.
- Billings, L. & C. Rudin. 1996. Optimality and Superiority: A new approach to multiple-wh ordering. In *Proceedings of FASL-3: The College Park meeting*, 35-60.
- Blaszczak, J. 2002. *Multiple WH-Konstruktionen im Slavischen*. Ms.: University of Potsdam, Germany.
- Blaszczak, J. & S. Fischer. 2001. Multiple wh-Konstruktionen im Slavischen: state of the art report. In *Linguistics in Potsdam 14*, 1-116. Potsdam: Univ.-Bibliothek, Publ.-Stelle.
- Bošković, Ž. 2001. On the interpretation of multiple questions. In *Linguistic Variation Yearbook 1*, 1-15.
- Bošković, Ž. 2000. What is special about multiple wh-fronting? In Hirotani et al. (eds.): *Proceeding of NELS 30*, Volume 1, 83-107. University of Massachusetts, Amherst.
- Bošković, Ž. 1998a. Multiple wh-fronting and Economy of Derivation. In *Proceedings of the 16th WCCFL* held at the University of Washington. Stanford: CSLI Publications.
- Bošković, Ž. 1998b. Wh-phrases and wh-movement in Slavic. In *Proceedings of Comparative Slavic Morphosyntax*, Bloomington, Indiana.
- Comorovski, I. 1996. *Interrogative Phrases and the Syntax-Semantics Interface*. Dordrecht: Kluwer.

- Franks, St. & T.H. King. 2000. *A Handbook of Slavic Clitics*. Oxford, Oxford University Press.
- Givón, T. 1976. Topic, Pronoun, and Grammatical Agreement. In Li (ed.): *Subject and Topic*, 149-88. New York & London: Academic Press Inc.
- Grewendorf, G. 2001. Multiple wh-Fronting. In *Linguistic Inquiry* 32(1), 87-122.
- Grohmann, K. K. 1998. Syntactic Inquiries into Discourse Restrictions on Multiple Interrogatives. In Abraham, (ed.) *Groninger Arbeiten zur germanistischen Linguistik* 42, 1-60.
- Guentchéva, Z. 1994. *Thématisation de l'objet en bulgare*. Bern: Peter Lang.
- Gundel, J. 1988. Universals of topic-comment structure. in Hammod et al. (eds.): *Studies in Syntactic Typology*, 209-39. Amsterdam: John Benjamins.
- Jaeger, T.F. 2003. Topics First! In- and Outside of Bulgarian Wh-Interrogatives. In *Proceedings of the HSPG 2003 Conference*, 181-202. Stanford: CSLI Publications.
- Jaeger, T.F. 2002. *On the Optional and Obligatory Realization of the Bulgarian Direct Object Clitic*. Ms. Stanford University.
- Jaeger, T.F. & V.A. Gerassimova. 2002. Bulgarian word order and the role of the direct object clitic in LFG. In Butt, M. & T.H. King (eds.): *Proceedings of the LFG02 Conference*. Stanford: CSLI Publications.
- Jaeger, T.F. & D.Y. Oshima. 2002. Towards a Dynamic Model of Topic Marking. In *Pre-proceedings of the Information Structure in Context Workshop*, 153-67. Stuttgart University.
- Kazenin, K. 2000. *On Coordination of Wh-Phrases in Russian*. Ms., Tübingen University.
- Keenan-Ochs, E. & B. Schieffelin. 1976. Foregrounding referents: A reconsideration of left-dislocation in discourse. In *Proceedings of BLS 2*, 240-57.
- Kuno, S. & K. Takami. 1993. *Grammar and discourse principles: Functional Syntax and GB theory*. Chicago: Chicago University Press.

- Lambova, M. 2003. On the Interaction of Multiple (Non) wh-Fronting and Multiple Topicalization in Bulgarian. In: *PennWPL* 8, 127-142.
- Lambrecht, K. 1994. *Information structure and sentence form. Topic, Focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- Leafgren, J.R. 1997. Bulgarian Clitic Doubling: Overt Topicality. In *Journal of Slavic Linguistics* 5(1), 117-43.
- Pavlov, I.V. 2000. *The non-argument functional field of Bulgarian clause structure*. Ph.D. thesis, NTNU Trondheim.
- Pesetsky, David. 1987. WH-in situ: Movement and Unselective Binding. In Meulen and Reuland (eds.): *The Representation of (In)definiteness*, 98-129. Cambridge, MA: MIT Press,.
- Rudin, C. 1997. AgrO and Bulgarian Pronominal Clitics. In *Proceedings of FASL: The Indiana meeting 1996*, 224-52. Michigan: Michigan Slavic Publications.
- Rudin, C. 1989. Multiple questions south, west and east: A Government-Binding approach to the typology of wh-movement in Slavic languages. In *International Journal of Slavic Linguistics and Poetics* 39-40.
- Rudin, C. 1988b. On multiple questions and multiple WH fronting. In *NLLT* 6, 445-502.
- Rudin, C. 1988a. Multiple Question in South Slavic, West Slavic and Romanian. In *Slavic and East European Journal* 32(1), 1-24.
- Rudin, C. 1985. *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*. Columbus, Ohio: Slavica Publishers.
- Reinhart, T. 1982. *Pragmatics and linguistics: an analysis of sentence topics*. Distributed by Indiana University Linguistics Club, Bloomington, Indiana.
- Richards, N. 1997. *What moves where in what language*. Ph.D. thesis, MIT.
- Scott, T. 2003. CP-Topic Parameter and Russian WH. Handout for *LSA Annual Meeting 2003*, Atlanta, GA.
- Steedman, M. 2000. Information Structure and the Syntax-Phonology Interface. In *Linguistic Inquiry* 31(4), 649-89.