
Configurationality and the Direct Object Clitic in Bulgarian*

VERONICA A. GERASSIMOVA

Stanford University
gerassim@stanford.edu

T. FLORIAN JAEGER

Stanford University
tiflo@stanford.edu

ABSTRACT - Bulgarian, as a free word order language without case, presents an interesting problem for the ongoing discussion whether or to what extent all languages are configurational (as defined in Nordlinger 1998:45). In this paper, we discuss the typological status of Bulgarian regarding its configurationality. First, we apply a set of configurationality tests (cf. King 1995, Speas 1990). Second, we show how the direct object clitic figures in determining possible constituent orders. Related to the direct object clitic, the role of information structure proves to be of crucial importance for the determination of word order in Bulgarian.

1 Introduction

In primarily rule-based grammar theories like GB and MP, it has generally been assumed that languages identify grammatical functions (GFs) by means of (GF)-configurationality, i.e. through phrase structure positions. However, it has been argued that, among others, languages like Walpiri (Hale 1981, Austin & Bresnan 1996, Simpson 1991), Wambaya (Nordlinger 1998) and Jiwari (Austin 1993) do not exhibit GF-configurationality. While the status of Walpiri is still debated (Legate *to appear*, Bresnan *class manuscript*, p.c.), the non-configurationality hypothesis has not been challenged for other languages.

Bulgarian presents a typologically interesting case for the discussion of configurationality. Unlike most other Slavic languages, Bulgarian has lost its case system. Only a small set of pronouns still shows overt case marking. Although traditionally considered to be underlyingly configurational (cf. Rudin 1985, 1990/1991, a.o.), spoken Bulgarian has a fairly free constituent order. SVO, SOV, VSO and marginally VOS and OVS constituent orders are possible with specific intonations.¹ Furthermore, Bulgarian allows direct and indirect object replication by an enclitic short-form pronoun. In this paper, we focus on the direct object clitic (DOC) and its impact on constituent order. In the case of direct object reduplication, all six theoretically possible constituents orders actually occur. Compare the contrast in acceptability in the following examples, depending on the presence of the 3rd person, plural direct object clitic *gi* 'them', which – if present – reduplicates the direct object NP *parite* 'them'.²

* The authors gratefully acknowledge the helpful support of the Californian sun and Moon Beans double Latte. We have benefited a lot from the suggestions of the three anonymous ESSLLI 2002 reviewers.

¹ Although crucial for the acceptability of the examples, prosodic information is omitted but not ignored here. Rudin (1990/1991:430), while completely ignoring intonation, judges all six word orders to be grammatical with different emphasis. However, the native speaker judgments of one of the authors (V.G.) and Avgustinova (1997) support the judgments presented here. We will come back to the importance of the intonation in section 3.

² We use the following glosses: 1, 2, 3 – first, second, and third person, DEF – definite article, FEM – feminine, M – masculine, PL – plural, refl – reflexive pronoun, SBJ – subjunctive marker, SG – singular.

S-(CL)-V-O (subject-clitic-verb-object)

- (1) Veronica (gi) vze parite.
 Veronica them took.3.SG money.DEF
Veronica took the money.

S-O-(CL)-V

- (2) Veronica parite (gi) vze.

V-(CL)-S-O

- (3) vze (gi) Veronica parite.

V-??(CL)-O-S

- (4) Vze ??(gi) parite Veronica.

O-??(CL)-V-S

- (5) Parite ??(gi) vze Veronica.

O-S*(CL)-V

- (6) Parite Veronica *(gi) vze.
 Intended: *Veronica took the money.*

To sum up, while Bulgarian cannot use case to identify GFs, it has very free word order. Thus the question arises how Bulgarian speakers identify GFs or, in other words, how GFs are encoded in this language. We will now examine whether there is any evidence for an underlying configurationality in Bulgarian.

2 How configurational is Bulgarian?

Traditionally (cf. Penčev 1973), and in more recent GB (cf. Dyer 1992, Rudin 1988, 1989, 1990/1991, Embick & Izvorski 1994) and MP approaches (cf. Rudin 1994, 1997, Embick & Izvorski 1998, Rudin et al. 1998, Rudin et al. 1999, Arnaudova 2001, Bošćovic 1998), GF-configurationality of Bulgarian has never been questioned, or provided evidence for but implicitly assumed.³ A number of configurationality tests have been proposed in the literature (e.g. Speas 1990:137, King 1995:Chapter 3). The problem with some of the tests is their theory-dependency. As we will see below, many of the test results have been argued to be explainable by means other than phrase structure configurationality. Here, we proceed as follows. We apply those tests which are applicable to Bulgarian (i.e. extraction, multiple *wh*-questions, binding, and weak crossover and other constituency tests, namely ellipsis, pronominalization, fronting and coordination) and discuss their outcome. We will see that, according to the results of these diagnostics, the GF-configurational status of Bulgarian is by no means as clear as commonly assumed.

One important characteristic of configurational languages is that they exhibit clear subject-object asymmetries such as Superiority effects, difference in possible coreference, and weak crossover, among others (Speas 1990:137). Such phenomena have been attributed to the distinct phrase structure positions of the subject and the object in configurational languages. Note that in the following we will sometimes speak about VP(-tests) when we mean (tests for) a verb-object constituent. We stick to this common terminology, although it is sometimes confusing.

A. In extraction, traces must be properly governed, according to the Empty Category Principle (ECP). The assumption is that objects are directly governed and assigned a theta-role by the verb in the VP. Thus objects extract freely, whereas subjects are governed indirectly, and therefore obey more extraction constraints (King 1995:54). In English, for example, objects but not subjects may extract from embedded clauses

³ For example, Dyer (1992:70) assumes the following c-precedence hierarchy for the basic word order in Bulgarian: SUBJ – V – O_{direct} – O_{indirect}. At the same time, he identifies forty-nine possible word orders for sentences with only a ditransitive verb and its complements (ibid:56).

with an overt complementizer. In Bulgarian, however, subjects and objects extract with equal ease. Consider the following examples of a noun modified by a relative clause, which involves extraction from an embedded clause with an overt complementizer:

Subject extraction

- (7) ženata, kojato Todor kaza, če _ e vidjala Marija, ...
 woman.DEF who T. said.3.SG that _ is seen.3.FEM.SG M.
The woman, who Todor said that has seen Maria.

Object extraction

- (8) ženata, kojato Todor kaza, če Marija e vidjala _, ...
 woman.DEF who T. said.3.SG thatM. is seen.3.FEM.SG _
The woman, who Todor said that Maria has seen.

Examples (7) and (8) are cases of subject and object relativization, respectively, and both examples are grammatical. The same lack of asymmetry in extraction between subjects and objects is also observed in *wh*-questions and topicalization. This leads Rudin (1985) to the conclusion that the ECP is irrelevant for Bulgarian. Thus extraction in Bulgarian fails to provide positive evidence for a structural difference between subjects and objects. According to the ECP, the identical behavior of subjects and objects can be attributed to the fact that both are directly governed by the verb, and are therefore inside the VP. In theory neutral terms, these data suggest that the subject and the object in Bulgarian have identical structural positions.

B. In multiple *wh*-questions in Bulgarian, all *wh*-words must be fronted, and they have been reported to exhibit strict ordering with subjects appearing first and the rest of the *wh*-words in specific positions (Rudin 1990/1991:440, 1988, 1989):

Wh-extraction according to Rudin (1990/1991, 1988, 1989)

- (9) Koj kogo vidja?
 who whom saw.3.SG
Who saw whom?
- (10) *Kogo koj vidja?
 whom who saw.3.SG
Whom did who see?

This subject-object asymmetry, too, has been accounted for in terms of phrase-structure and the Superiority condition (cf. Chomsky 1973). However, in his analysis of multiple *wh*-questions, Boškovic (1993, 1998) argues that, in Bulgarian, only the first extracted *wh*-word is subject to the Superiority condition. Since the following *wh*-words occur in free order with respect to each other, they are unaffected by Superiority. Consider the following data taken from Boškovic (1993, cf. 1998:60-61):

Superiority and multiple wh-questions in Boškovic (1998:60-61, 1993)

- (11) Kogo kak e celunal Ivan.
 whom how is kissed Ivan
How did Ivan kiss whom?
- (12) *Kak kogo e celunal Ivan
 How whom is kissed Ivan
- (13) Koj kogo kak e tselunal
 Who whom how is kissed
Who kissed whom how?
- (14) Koj kak kogo e celunal
 Who how whom is kissed
Who kissed whom how?

Boškovic (1998:62, 1993) concludes that Superiority can only account for the first fronted *wh*-word in multiple *wh*-questions and argues that Superiority can only be understood as a descriptive generalization. An alternative explanation of these ordering effects can be attributed to a typologically motivated GF hierarchy constraint. Restricted

to the first *wh*-expression in the string, it predicts that the least oblique constituents will precede all the rest (cf. Billings & Rudin 1994 for an OT account similar in spirit). Although the question requires further investigation, alternative data suggest that animacy is one factor in the ordering of *wh*-words. (15) and (16) show that when the object but not the subject is of type [+animate], both orders are acceptable. If the [+animate] object precedes the [-animate] subject, as in (16), the DOC is obligatory. (17) indicates that a DOC does not repair the ungrammaticality of (10). Therefore the DOC cannot be the direct cause of the constituent reordering in (16) but is better thought of as an indirect consequence of the *wh*-word reordering, which in turn is licensed by a semantic animacy hierarchy.

Wh-extraction with reversed animacy of object and subject

- (15) Kakvo kogo iznenada?
 what whom surprised.3.SG
What surprised whom?
- (16) Kogo kakvo go iznenada?
 whom what 3.MASC.SG surprised.3.SG
What surprised whom?
- (17) *Kogo koj go vidja
 whom who him saw.3.SG
Whom did who see?

Since examples like (16) violate Superiority, we conclude that in Bulgarian the ordering of *wh*-elements in multiple *wh*-questions cannot be explained by Superiority. Thus multiple *wh*-questions do not provide clear evidence about the configurationality of Bulgarian.

C. It has been proposed in the literature that co-reference relations between a lexical NP and a pronoun differ between subjects and objects. This is accounted for by Principle C of the GB Binding Theory (cf. Chomsky 1981), which states that R-expressions (e.g. lexical NPs and *wh*-traces) must be free. That is, they cannot be co-indexed with a c-commanding antecedent. For configurational languages, in which the subject, by definition, c-commands the object, Principle C predicts subject-object asymmetries for binding (e.g. Speas 1990:137, King 1995:50). Examples (18) and (19) illustrate co-reference relations between a pronominal argument and a lexical NP embedded in a relative clause:

Principle C effects on anaphoric binding

- (18) Kotkata, s kojato Todor_i živeeše, go_i napusna.
 cat.DEF with whom.FEM.SG T. lived.3.SG him left.3.SG
The cat who Todor_i lived with left him_i.
- (19) *Toj_i napusna kotkata, s kojato Todor_i živeeše.
 he left.3.SG cat.DEF with whom.FEM.SG T lived.3.SG
**He_i left the cat who Todor_i lived with.*

In (18) the pronoun go is the object of the main clause, and can be coreferential with Todor. In (19), however, the pronoun is the subject of the main clause, and a co-referent reading is impossible. Following Principle C, the contrast between (18) and (19) can be explained in terms of the different structural positions of go and Todor: in (18), the lexical NP and the pronoun do not c-command each other, whereas in (19) the pronoun is the subject, and therefore c-commands Todor. Consider the above sentences with the order of the lexical NP and the pronoun reversed, as in (20) and (21):

Precedence effects

- (20) *Napusna go_i kotkata, s kojato Todor_i živeeše.
left.3.SG him cat.DEF with whom.FEM.SG T. lived.3.SG
The cat who Todor_i lived with left him_i.
- (21) Kotkata_k, s kojato Todor_i živeeše, toj_i ja_k napusna.
cat.DEF with whom.FEM.SG T. lived.3.SG he her left.3.SG
He_i left the cat who Todor_i lived with.

In the parallel Russian examples, King (1995:52) assumes that in examples like (20) the subject NP is right-adjoined to VP, and in (21) the object is left-adjoined to IP. As a result, the pronoun and the lexical NP do not c-command each other, and the contrast in grammaticality is due entirely to their linear ordering. While this analysis is valid for Bulgarian in (21), in (20) the facts are more complex since the object pronoun is realized as a clitic, and not as a free pronoun. We follow Rudin (1997:238) in that go in fact commands⁴ Todor, which accounts for the ungrammaticality of (20).

However, these data can be accounted for entirely in terms of linear precedence: in both grammatical examples above go does not precede Todor, while in the ungrammatical ones the pronoun precedes the lexical NP. This follows from the fact that in the typically assumed right-branching structure dominance directly correlates with linear precedence. Thus from these data it is not clear whether c-command is a necessary condition on binding in Bulgarian, or whether the relevant generalizations have to do with linear and syntactic prominence, as proposed by Bresnan (1994). The discussion of weak crossover in the next section provides a more definite answer to this question.

D. Weak crossover is another standard test for configurationality. It is based on the observation that a pronoun cannot appear on the path between a trace and its operator. One explanation for this fact in English is the Leftness Condition, which says that a pronoun cannot be coindexed with a variable to its right (Chomsky 1976). It rules out cases like (22), in which a pronoun is coindexed with a *wh*-trace to its right, and licenses cases like (23), in which the *wh*-trace is to the left of the pronoun:

Weak crossover in English

- (22) *Who does his mother love?
who_i does his_i mother love t_i
- (23) Who loves his mother?
who_i t_i loves his_i mother?

In Bulgarian, the weak crossover data are slightly more complex. Like Russian (cf. King 1995:55) and unlike English, Bulgarian has reflexive and non-reflexive possessive pronouns. Reflexive possessives must be bound in their minimal domain. (24) is ungrammatical regardless of whether the reflexive is co-indexed with the object *wh*-word or not. In contrast, (25) is grammatical if and only if the reflexive is coreferential with the subject *wh*-word. Such examples show that reflexive possessives do not pattern together with the English possessives, and cannot be used to test crossover.

Reflexive possessives

- (24) *Kogo_i običa majka si_{i/j}?
Whom love.3.SG mother his-refl
Whom_i does his_i mother love?
- (25) Koj_i običa majka si_{i/*j}?
Who love.3.SG mother his-refl
Who_i loves his_i mother?

⁴ Whereas Rudin (1997) assumes that go is the head of AgrO, Jaeger & Gerassimova (*to appear*) use an LFG-account where go does not c-command but f-commands (a command notion that does not rely on phrase-structure, cf. Bresnan 2001) Todor. Our account does not draw on assumptions about configurationality.

Unlike reflexive possessives, non-reflexive possessives cannot be bound in their minimal domain. Because of that, example (26), the counterpart of the English (23), is ungrammatical under coreference..

Non-reflexive possessives in subject *wh*-questions

- (26) Koj_i običa majka mu_{*i/j}?
 Who love.3.SG mother his
Who_i loves his_i mother?

Some evidence relating to weak crossover comes from non-subject *wh*-questions like (27), corresponding to (25) above. Although the preferred reading for it is non-coreference between *kogo* and the possessive pronoun, co-indexation is also possible. In both cases the sentence is good, showing lack of crossover effects:

Non-reflexive possessives in object *wh*-questions

- (27) Kogo_i običa majka mu_{i/j}?
 Whom love.3.SG mother his
Whom_i does his_i mother love?

The declarative variants of the object *wh*-question are shown in (28) and (29), with different word orders. Note that (28) is ungrammatical in the case of coreference, although there are no binding violations. The reason for this is that the pronoun precedes the lexical NP. In contrast, (29), in which the pronoun neither precedes nor c-commands the antecedent, is grammatical⁵:

Non-reflexive possessives in declarative sentences

- (28) Majka mu_{*i/j} običa Ivan_i.
 mother his love.3.SG I.
*His_{*i/j} mother loves Ivan_i.*
- (29) Ivan_i go_i običa majka mu_{i/j}.
 I. him love.3.SG mother his
His_{i/j} mother loves Ivan_i.

In sum, the grammaticality of (27), which is the counterpart of the ungrammatical English (22), shows that, unlike English, Bulgarian lacks weak crossover. Instead, the data in this section can be explained entirely in terms of the conditions on Binding in Bulgarian: a pronoun may not precede its antecedent. This suggests that weak crossover can be better explained in terms of linear and/or syntactic prominence, as proposed in Bresnan (1994), Dalrymple et al. (2001).

In conclusion, we have established that subject-object asymmetry tests, such as Superiority, binding and weak crossover, seem to suggest that Bulgarian is not GF-configurational. Next, we apply constituency tests such as VP-pronominalization, fronting, and coordination (cf. Speas 1990) to Bulgarian.

E. Bulgarian lacks VP-pronominalization. In other words, it is impossible to use an overt anaphoric element to refer to a preceding verb phrase.⁶ Concerning VP-fronting, Bulgarian has no construction resembling it, either:⁷

⁵ The presence of the direct object clitic in (29) is due to obligatory clitic doubling of topicalized objects (cf. Dimitrova-Vulchanova & Hellan 1998).

⁶ Note, however, that VP-ellipsis presents an interesting question. In contrast to English, it is impossible to gap a VP to the exclusion of the auxiliary in Bulgarian whereas it is possible to gap VPs which are not governed by an I.

⁷ Examples like the following show that in some cases it is possible to front a subjunctive verb and its object. However, the fronted constituent is not the matrix VP but an subjunctive argument of *trudno* with a controlled subject:

[Da vzeme izpita] šte e trudno za Ivan
 SBJ take.3.SG the-exam will is difficult for I.
To take the exam will be hard for Ivan.

VP-fronting

- (30) *Ivan obešta, če šte vzeme izpita,
 I. promised.3.SG that will take.3.SG exam.DEF
 i da go vzeme toj šte.
 and SBJ it take.3.SG he will
Intended: Ivan promised to pass the exam, and pass it he did.

F. Finally, there is only one test, namely VP-coordination, which can be interpreted as evidence for the existence of a verb-object constituent in Bulgarian. Coordination tests are based on the observation that if two items can be coordinated, they are both constituents (Sag et al. 1985, Peterson 1981). As (31) shows, verb-object phrases can indeed be coordinated in Bulgarian, however (32) indicates that subject-verb phrases may also be conjoined:

Coordination of verb-object and verb-subject: [] indicates the coordination structure

- (31) Cjal den štjax [da piša pisma i da četa knigi].
 whole day would.1.SG SBJ write.1.SG letters and SBJ read.1.SG books
All day, I would write letters and read books.
- (32) [Az obiçam, a ti mrziš] sladoled.
 I love.1.SG and you hate.2.PL ice-cream
I love, and you hate ice-cream.

The contrast between (31) and (32) could be taken as an indicator that objects and verbs form a constituent together, which the subject is not a part of.

To sum up, we have seen that Bulgarian does not show conclusive evidence for GF-configurationality (i.e. the majority of the tests for subject-object asymmetries suggest that Bulgarian has a VP that contains subject and object at the same phrase structure level)⁸. Theories of grammar which are based on the assumption of a Universal Grammar (UG) specific to natural language (cf. Chomsky 1995, 1985 for the MP and GB respectively) are basically forced (or at least biased) to analyze all languages as configurational. That is, according to those theories, since there are languages which identify GFs by means of phrase structure (e.g. English) and since all languages are alike, all languages must be at least underlyingly GF-configurational.⁹ It is this effect of the theoretical framework on the interpretation of the empirical facts that is the point of departure for the next section.

3 The direct object clitic (DOC)

According to Nordlinger's (1998:49) typology of configurationality we would expect relatively strong head- or dependent-marking for a supposedly non GF-configurational language like Bulgarian. However, since the direct object clitic (DOC) only optionally reduplicates object NPs (cf. exx. (1)-(6), p. 2), Bulgarian is not a typical example of a strongly head-marking language.¹⁰ Furthermore, the lack of almost any case morphology (cf. section 1) leaves no doubt that Bulgarian is not dependent-marking.

⁸ It remains open for further research to investigate how the apparent mismatch of the outcome of the different tests can be resolved. Here we cannot provide a full conclusive account of the data.

⁹ We are very well aware that relations, such as MOVE and AGREE, between the universal deep-structure and the specific surface realization of it can be seen as metaphors of the theory, so that the absence of configurationality does not disprove a program like the MP. Here we only want to raise the point that for languages like Bulgarian, Jiwari and others there seems to be no evidences for configurationality if we do not assume a UG.

¹⁰ See Helmbrecht (2001) on head-marking vs. dependent-marking languages.

Configurationality and the Direct Object Clitic in Bulgarian

This fact, puzzling at first glance, is not at all surprising if we take into consideration information structure (cf. Vallduvi 1993, 1992, Lambrecht 1994). As shown in Jaeger & Gerassimova (*to appear*), the encoding of information structure through intonation and the DOC has a crucial effect on word order in Bulgarian. This in turn means that information structure indirectly determines not only the assignment of discourse functions (e.g. TOPIC and FOCUS; henceforth DFs), but also the identification of GFs. Bulgarian thus groups with other languages, such as Hungarian (cf. Kiss 1987), which are commonly called discourse-configurational (cf. Kiss 2001, 1995, Nordlinger 1998). Other authors (e.g. Rudin 1997, 1994, Rudin et al 1998, 1999, Embick & Izvorski 1994, 1997) have provided extensive evidence for the DF-configurationality of Bulgarian, more precisely its left-periphery configuration, which we cannot go into here due to the lack of space.

In this section, we show that word order in spoken Bulgarian depends heavily on the DOC.¹¹ Consider the following data based on Avgustinova (1997:112, examples with adverbs have been added by us).¹² The question-answer pairs were elicited from native speakers. Before each question, which determines the focus of the answer, another sentence introduces the topic (i.e. the most salient part of the ground), as in 'What about X_{topic}? What_{focus} did X_{topic} give_{tail} to Y_{tail}?':

Topic	Context	Elicited answer in Bulgarian
<i>Andrej</i> (Andrew)	What did he do?	1. Andrej [izjade boba].
	What did he eat?	2. Andrej [BOBA] izjade. 3. Andrej izjade [BOBA]
	What did he do to the beans?	4. Andrej [go IZJADE] boba. 5. Andrej boba [go IZJADE]. 6. [IZJADE go] boba Andrej.
	When did he eat the beans?	7. Andrej izjade boba [VČERA].
<i>boba</i> (the beans)	What happened to them?	8. Boba [go izjade Andrej].
	What did Andrew do to them?	9. Boba [go IZJADE] Andrej. 10. Boba Andrej [go IZJADE]. 11. [IZJADE go] Andrej boba.
	Who ate them yesterday?	12. Boba [ANDREJ] go izjade. 13. Boba go izjade [ANDREJ]. 14. [ANDREJ] (go) izjade boba.
	When did Andrew eat them?	15. Boba [VČERA] go izjade Andrej. 16. Boba Andrej go izjade [VČERA]
<i>včera</i> (yesterday)	What happened then?	17. Včera [Andrej izjade boba].
	What did Andrew do then?	18. Včera Andrej [IZJADE BOBA].
	What did Andrew eat then?	19. Včera Andrej izjade [BOBA].
	Who ate the beans then?	20. Včera boba go izjade [Andrej].
<i>izjade</i> (eat)	I know that Andrew cooked but what did he eat?	21. Izjade Andrej [BOBA].
	I know who cooked the beans but who ate them?	22. [ANDREJ] izjade boba. 23. Izjade go boba [ANDREJ].

Table 1

¹¹ As the tables below indicate intonation is also of crucial importance for the understanding of Bulgarian word order. However, due to limited space we focus on the function of the DOC here.

¹² In each example, the phonologically most prominent element of the TOPIC (i.e. the LINK, cf. Vallduvi 1993, 1992) is printed bold. Small caps indicate emphatic stress. FOCUS is marked by []. In all examples, *Andrej* (Andrew) is the subject, *boba* (the beans) the object and *izjade* (eat-PERFECTIVE-3.SG.[SUBJ]) the predicate. *Včera* (yesterday) is used to show where adjuncts can occur. *Go* is the 3.SG.MASC pronominal direct object clitic. It is coreferent with the direct object NP *boba* in all examples.

In all of the above examples the DOC agrees with the object NP in person, number, and gender and thereby identifies the object. Together with the verbal subject-agreement morphology and prosody the DOC provides important cues for the hearer to identify the GFs in a sentence. We will now show that the DOC provides cues about the information structure of the sentence as well.

TABLE 2 shows – for all six theoretically possible constituent orders of a sentence with a transitive predicate - how constituent order and the assignment of GFs and DFs interact (cf. TABLE 1). We can read each row in TABLE 2 as a set of conditions that have to be satisfied to allow that word order. For a given constituent order schema (COS), realization of the object clitic is indicated by CL, optionality of the object clitic by (), emphatic stress by EM, and "-" for the absence of any of the above. _F indicates FOCUS, and _T TOPIC. The data are classified with respect to information packaging using Vallduvi's (1992, 1993) framework. GROUND, as the old information, is distinguished from FOCUS, (i.e. the new information) and further divided into TOPIC (i.e. what the information provided in the sentence is about) and TAIL (i.e. the rest of the background information). The fifth column refers to the examples from TABLE 1 which have the relevant word order.

COS	SUBJ	VERB	OBJ	Ex.
C1	- _T /-	- _F	- _F	1, 18
C2	- _F	- _F	- _F	17
C3	- _T /-	-	EM _F	3, 19
C4	- _T	-	-	7
C5	EM _F	- _T	-	22
C6	EM _F	-/(CL)	- _T	14
C7	- _T	EM _F /CL	-	4
COS	SUBJ	OBJ	VERB	Ex.
C8	- _T	EM _F	-	2
C9	- _T	-	EM _F /CL	5
COS	VERB	SUBJ	OBJ	Ex.
C10	EM _F /CL	-	- _T	11
C11	- _T	-	EM _F	21

COS	VERB	OBJ	SUBJ	Ex.
C12	EM _F /CL	-	- _T	6
C13	-/CL	- _T	EM _F	23
COS	OBJ	SUBJ	VERB	Ex.
C14	- _T	-	EM _F /CL	10
C15	- _T	-	-/CL	16
C16	- _T	EM _F	-/CL	12
COS	OBJ	VERB	SUBJ	Ex.
C17	- _T	- _F /CL	- _F	8
C18	- _T	EM _F /CL	-	9
C19	- _T	-/CL	-	15
C20	- _T /-	-/CL	EM _F	13, 20

Table 2

So far, we are not entirely clear what the exact function of the DOC is but it certainly marks a discourse function. However, it is a striking fact that when the DOC is present, either the object is the topic¹³ or the verb receives strong stress (cf. TABLE 2). Since the obligatory presence of the DOC in the case of a stressed verb may be due to phonological constraints, we ignore those cases for now. The DOC can reduplicate the object NP if and only if the object is assigned the information structural function of a topic.¹⁴ Because the DOC is coreferent with the reduplicated NP, it identifies the object-GF as well as the topic-DF. In other words, the DOC takes over some of the functions

¹³ For a detailed discussion of the DF-marking properties of the DOC and an analysis what kind of TOPIC-DF the DOC marks, we refer the interested reader to Jaeger & Gerassimova (*to appear*).

¹⁴ This analysis is supported by observations made independently from our research (cf. Leafgren's 1997a, 1997b). Note that regardless of the exact definition of TOPIC, the fact that the DF marked by the DOC correlates with old information (cf. Prince 1981) indirectly accounts for the fact that only specific object NPs can be reduplicated by the DOC (cf. Dimitrova-Vulchanova & Hellan 1998:xvii). We are grateful to the anonymous reviewer, who reminded us of the importance of specificity on direct object reduplication in Bulgarian.

performed by phrase-structure configurationality in other languages (e.g. English). It identifies a grammatical function.

To sum up, in previous work it has been established that Bulgarian shows a certain degree of DF-configurationality. We have shown that word order in Bulgarian indeed depends more on information structure than on phrase-structural GF-assignment. The striking feature of Bulgarian is that, instead of using phrase structure to identify GFs, the primary means of DF and GF identification are prosody, lexical markers (e.g. the DOC) and morphology (e.g. subject-verb agreement).

4 Conclusions

We have shown that Bulgarian by no means provides conclusive evidence for GF-configurationality. This posits a strong empirical problem for most existing accounts of Bulgarian (especially those within the GB and MP frameworks). In a second step, we have argued that, for spoken Bulgarian, prosody and the DOC (in interaction with the subject-verb agreement morphology) provides alternative means for GF- and (before all) DF-assignment. We have investigated one of those alternative means, the DOC, in more detail and presented data that suggest how the DOC figures in the identification of DFs and GFs. Further research and a formal implementation of our observations within Lexical Functional Grammar (cf. Bresnan 2001) is presented in our ongoing work (cf. Jaeger & Gerassimova *to appear*).

References

ARNAUDOVA, OLGA

- (2001): *Prosodic movement and information focus in Bulgarian*; in: *Proceedings of Formal Approaches to Slavic Linguistics 9*; Michigan Slavic Publications.

AUSTIN, PETER

- (1993): *Word order in free word order languages: the case of Jiwarli*; in: *Language*.

AUSTIN, PETER / BRESNAN, JOAN

- (1996): *Non-configurationality in Australian Aboriginal languages*; in: *Natural Language and Linguistic Theory 14* (2), p. 215-268.

AVGUSTINOVA, TANIA

- (1997): *Word Order and Clitics in Bulgarian*. in: *Saarbrücken Dissertations in Computational Linguistics and Language Technology, Volume 5*. DFKI, Saarbrücken.

BAKER, MARK C.

- (1991): *On some subject/object asymmetries in Mohawk*; in: *Natural Language and Linguistic Theory 9*, p. 537-576.

BRESNAN, JOAN

- (2001): *Lexical Functional Syntax*; Blackwell Publishing, Cambridge.
- (1994): *Linear Order vs. Syntactic Rank: evidence from weak crossover*; in: *Proceedings of the Chicago Linguistics Society 30*, p. 57-89.
- (1982): *The Mental presentation of Grammatical Relations*; Cambridge, MIT Press.

BILGINS, LOREN / RUDIN, CATHERINE

- (1994):

CHOMSKY, NOAM

- (1995): *The Minimalist Program*; Cambridge, MIT Press.
- (1986): *Conditions on Rules of Grammar*; in: *Linguistic Analysis 2*, p. 303-350.
- (1981): *Lectures on Government and Binding*; Dordrecht, Foris.

DALRYMPLE, MARY, RONALD KAPLAN, AND TRACY KING

- (2001): *Weak Crossover and the Absence of Traces*. A paper presented at LFG2001.

DIMITROVA-VULCHANOVA, MILA / HELLAN, LARS

- (1998): *Introduction*; in Dimitrova-Vulchanova, M. / Hellan, L. (eds.): *Topics in South Slavic Syntax and Semantics*, pp. ix-xxvii; Amsterdam & Philadelphia, John Benjamins Publishing.

DYER, DONALD L.

- (1992): *Word Order in the simple Bulgarian Sentence: A Study in Grammar, Semantics and Pragmatics*; Amsterdam – Atlanta, Rodopi B. V.

EMBICK, DAVID / IZVORSKI, ROUMYANA

- (1997): *Participle-Auxiliary Word Order in Slavic*; in: *Annual workshop on formal approaches to Slavic linguistics. The Cornell meeting 1995*, pp. 210-239; Michigan, Michigan Slavic Publications.
- (1994): *On long head movement in Bulgarian*; in: *Proceedings of the Eleventh Eastern States Conference on Linguistics, 1994*, pp. 104-115; Ithaca, Cornell University.

HALE, KENNETH

- (1981): *On the position of Walpiri in a typology of the base*; distributed by Indiana University Linguistics Club, Bloomington.

HELMBRECHT, JOHANNES

- (2001): *Head-marking vs. dependent-marking languages*; in Haspelmath, M. / Koenig, E. / Oesterreicher, W. / Raibler, W. (eds.): *Language Typology and Language Universals*, pp. 1424-1432; Berlin & New York, de Gruyter.

JAEGER, T. FLORIAN / GERASSIMOVA, VERONICA

- (to appear): *Bulgarian word order and the role of the direct object clitic in LFG*; to appear as electronic publication: *Proceedings of the LFG 2002 Conference*; Stanford, CSLI Publications.

KING, TRACEY H.

- (1995): *Configuring Topic and Focus in Russian*. in: *Dissertations in Linguistics*, CSLI Publications, Stanford.

KISS, KATALIN É.

- (2001): *Discourse configurationality*; in Haspelmath, M. / Koenig, E. / Oesterreicher, W. / Raibler, W. (eds.): *Language Typology and Language Universals*, pp. 1442-1455; Berlin & New York, de Gruyter.
- (1995): *Introduction*; in Kiss, K. E. (ed.): *Discourse Configurational Languages*, pp. 3-27; Oxford, Oxford University Press
- (1987): *Configurationality in Hungarian*; Dordrecht, Boston, D. Reidel Publishers.

LAMBRECHT, KNUD

- (1994): *Information structure and sentence form. Topic, Focus, and the mental representations of discourse referents*; Cambridge, Cambridge University Press.

LEAFGREN, JOHN R.

- (1997b): *Definiteness, Givenness, Topicality and Bulgarian Object Reduplication*; in: *Balkanistica 10*, pp. 296-311.
- (1997a): *Bulgarian Clitic Doubling: Overt Topicality*; in: *Journal of Slavic Linguistics, Volume 5, Number 1*, pp. 117-143.

LEGATE, JULIE A.

- (2002): *The configurational structure of a nonconfigurational language*; in: *Linguistics Variation Yearbook*.

NORDLINGER, RACHEL

- (1998): *Constructive Case: Evidence from Australian languages*; in: *Dissertations in Linguistics*; Stanford, CSLI Publications.

PETERSON, P.

- (1981) *Problems with Constraints in Coordination*; in: *Linguistic Analysis No. 8*, pp. 449-460.

PRINCE, ELLEN

- (1981): *Toward a Taxonomy of Given-New Information*; in Cole, Peter (ed.): *Radical Pragmatics*; e.g. New York, London, Academic Press.

RUDIN, CATHERINE

- (1997): *AgrO and Bulgarian Pronominal Clitics*; in: *Annual workshop on formal approaches to Slavic linguistics. The Indiana meeting 1996*, p. 224-252; Michigan, Michigan Slavic Publications.
- (1994): *On focus position and focus marking in Bulgarian questions*; in: *Papers from the Fourth annual meeting of the Formal Linguistic Society of Midamerica, April 15-18, 1993*, p. 252-266, Iowa City.
- (1990/1991): *Topic and Focus in Bulgarian*; in: *Acta Linguistica Hungarica, Vol. 40 (3-4)*, p. 429-449.
- (1988): *On multiple questions and multiple WH fronting*; in: *Natural Language and Linguistic Theory 6*, p. 445-502.
- (1985): *Aspects of Bulgarian Syntax: Complementizers and WH Constructions*; Columbus, Ohio: Slavica Publishers.

RUDIN, CATHERINE / KING, TRACY HOLLOWAY / IZVORSKI, ROUMYANA

- (1998): *Focus in Bulgarian and Russian yes-no questions*; in: *University of Massachusetts Occasional Papers 21. Proceedings of Workshop on Focus*, p. 209-226, Amherst.

SAG, I., G. GAZDAR, T. WASOW, AND S. WECHSLER

- (1985): *Coordination and How to Distinguish Categories*; in *NLLT No. 3*, pp.117-172.

SIMPSON, JANE HELEN

- (1991): *Aspects of Walpiri Morphology and Syntax*; Doctoral Dissertation, MIT, Cambridge.

SPEAS, MARGARET

- (1990): *Phrase Structure in Natural Language*; Kluwer Academic Publishers, Dordrecht.

VALLDUVI, ENRIC

- (1993): *Information Packaging: A Survey*; MS. Centre for Cognitive Science and Human Communication Research Centre, University of Edinburgh.
- (1992): *The Informational Component*; New York, Garland.