Saliency and Contrast in Colloquial Bulgarian: Clitic Left Dislocation versus Contrastive Topicalization*

Olga Arnaudova
University of Ottawa
arnaudova_olga@yahoo.com

Abstract
The claim advanced in this paper is that the presence of a left-dislocated element together with a resumptive clitic in Bulgarian is a special case of argument saturation with implications for the focus structure of the clause, while contrast involves discontinuous focus (contrastive topics/foci) with no clitics present in the derivation. Contrastive topic/focus constructions in Bulgarian can be united on the view that they involve (sets of) ordered pairs where the higher element is valuing a contrastive feature (cf. OCC in Chomsky 2001) while the element in the VP is a non-contrastive topic or focus. The contrastive feature participates in wh-structures but not in clitic-left-dislocated structures where pairing between arguments is ‘accidental’.

1 Introduction
In this paper, I discuss two distinct types of topics: inherent topics in clitic-left-dislocated structures above the CP domain, and topics participating in topic-focus sets in the TP and vP domain of the clause. The analysis draws on the view that the predicate-argument relation is not necessarily obtained uniquely within the VP(vP) and that Clitic Left Dislocation Structures ‘externalize’ an argument by saturating it in a higher domain.

The paper is organized in the following way. First, in section 2, I present Clitic Left Dislocation in Bulgarian, a construction which until recently has been overlooked in generative studies of Bulgarian. It involves obligatory clitic resumption and thematic ‘redundancy’. Section 3 contains a general overview of clitics, their relation to focus, and syntactic properties. Briefly summarizing some previous accounts, I adopt the view that clitics are

* Parts of this paper, which develops ideas from chapter 8 of my doctoral thesis, were presented at the Dislocated Elements Workshop, ZAS Berlin (November, 2003), FASL13, The Linguistics Mini-conference, University of Ottawa (May, 2004), and the Meeting of the Canadian Linguistic Association in Winnipeg, Manitoba (May, 2004). The author is grateful for comments and feedback received during those meetings and to the editors of this volume for their helpful suggestions. This research is supported by SSHRC Research Grant # 410-2003-0167 to María Luisa Rivero. All mistakes are my own.
argument variables generated in Specifier positions in the verbal domain. Section 4.1 deals with contrastive topics, which are never doubled by clitics, and their semantics. Contrastive topics are argued to be semantically related to (sets of) ordered pairs and to value a contrastive feature, while CLLDed topics are inherently topical in nature and are not ordered. They do not depend for their interpretation on a focused item found in the clause and are never focused themselves. Section 5 discusses T-F ‘pairings’ and the fact that CT-F (and Contrastive Focus-Topic) relations can be obtained in a ‘split focus’ fashion where the higher element of the pair is related to a set of alternatives and the lower element is non-contrastive. As shown in 5.3, contrastive (focus) features are not limited to declaratives and are also found in wh-questions. Section 6 provides additional evidence that CLLD topics are quite distinct from contrastive topics and are situated in a domain above the TP-level containing contrastive topic/focus. In section 7.1, I present Baker’s Polysynthetic Parameter, according to which arguments are basically generated in two predication domains. The general proposal for Bulgarian is outlined in section 7.2. I propose that CLLD and contrastive structures in Bulgarian exemplify two different types of argument saturation. In the case of CLLD, saturation is achieved through clitic variables but is incomplete; consequently, a ‘double’ can properly saturate the predicate. Conclusions are given in section 8.

2 Clitic Left Dislocation Structures

It has been noted for Bulgarian that in some cases a noun phrase or a full pronoun can ‘double’ a clitic in the same sentence. These sentences seem to have ‘redundant’ thematic structure, since the verb appears to be simultaneously assigned two identical thematic roles: one to the clitic, and one to the noun. Clitic reduplication/doubling, in general, is cited as one of the distinguishing characteristics of Bulgarian (see Franks and King 2000 and Rudin 1991 for references and discussion) but has been described as ‘optional’. As proposed in Arnaudova 1999, 2002, these constructions are cases of clitic left dislocation rather than doubling constructions.² CLLD has

¹ According to Rudin (Rudin 1991, n. 10), ‘clitic doubling is obligatory in some idioms and highly preferred in a few other constructions but is generally optional.’ Franks and King (2000: 251) note that ‘in Bulgarian, most clitic doubling is, at least superficially, optional.’ In representative grammars of Bulgarian, it is acknowledged that we may have to deal with two different varieties of Bulgarian and with some kind of ‘avoidance’ of certain structures in literary Bulgarian due to influence from Russian (see Andrejcin et. al 1977: §522, p. 376).

² Clitic doubling is a case where the clitic and the noun, in situ, are both ‘arguments’, as in the Spanish Lo vimos a Juan. The noun in this case is interpreted as focused information. Clitic Left Dislocation was studied first by Cinque (1990) for Italian — a language where clitic doubling is not attested — and in many other languages, such as Greek (Iatridou 1991)
no syntactic analogue in English or in any other Slavic language. In (1), a left-dislocated element is obligatorily linked to a coindexed resumptive clitic pronoun.

(1) Ivan Marija go vidja.
\[ \text{Ivan } \text{Mary } \text{Cl-ACC} \text{ see-Past} \]
‘Mary has seen Ivan.’

In (2), multiple nouns appear as dislocated, each doubled by a clitic (with the exception of dislocated subjects, such as az ‘I’, which do not have a corresponding (overt) clitic:

(2) Az uchebnika na Stojan *(mu) *(go) dadox.
\[ I \text{ textbook-the to Stojan Cl-DAT Cl-ACC give-1P,Past,Sg} \]
‘I gave to Stojan the textbook.’
(lit. ‘As for me, the textbook and Stojan, (I) gave it to him.’)

Clitic left-dislocated elements are not limited to DPs and can also be CPs. An example with a CLLDed CP is given in (3):

(3) [CP Che Simeon shte specheli izborite] go znajat vsichki.
\[ that Simeon \text{ will win elections-the Cl-ACC.SG know everybody} \]
‘Everybody knows that Simeon will win the elections.’

The phenomenon of clitic reduplication in constructions where a left-dislocated noun appears in the clause is attributed to the specific, definite, topical, or referential features of the nominal expressions (see, e.g., Guentcheva 1994, where it is argued that ‘clitic doubling’ codes the ‘themacity’ of the object). In other studies, it is argued that clitics are realization of (optional) agreement markers associated with functional agreement projections (Franks and King 2000) and, finally, clitics are seen as arguments of the verb (Penchev 1993). Therefore it is worth exploring if clitics are indeed topicality markers and what their status in grammar is.

3 Bulgarian Clitics as Argument Variables

Bulgarian clitics are non-tonic, short non-emphatic forms of the personal pronouns of the first, second and third person singular and plural, encoding features of the direct object (go-CL.Sg Masc; ja-CL, Sg, Fem, gi-CL,Pl), the indirect object (mu-CL.Sg, Masc, j-CL, Sg, Fem; im-CL, Pl) and, in noun

and French (Hirschbühler 1975). Alexiadou & Anagnostopoulou (1998) propose for Greek that the subject does not occur in the EPP position but in the thematic vP domain or in an A’-adjoined position coindexed with a null argumental pronominal. See also Krapova 2002 for Bulgarian. Krapova & Cinque (2003) have observed that CLLD is also attested in wh-questions.
phrases, of the possessor. Every clitic has a corresponding full pronoun form (go – nego; ja – neja; gi – tjax; mu – na nego; j – na neja; im – na tjax). There are no subject clitics in Bulgarian but recently SE has been claimed to be an overt manifestation of subject-like properties (see Rivero 2001 for more discussion). Clitics never co-exist with full forms of the personal pronouns unless focus is involved.3 In this section I will present some evidence for this claim and for their argument-like behaviour.

The relative order of clitics is strictly indirect object–direct object, as in (4a), whereas the order in the presence of a full pronoun is reversed, as in (4b), and two full pronouns cannot be used to replace both clitics regardless of the order, as shown in (4c):

(4) a. Dadox mu gi.
give-1P,Past,Sg-1St Cl-DAT.MASC Cl-Pl.
‘I gave them to him.’

b. Dadox gi na nego/na Ivan.
give-1P,Past,Sg Cl-Pl to him/to Ivan

c. *Dadox tjax na nego/*Dadox na nego tjax.
gave-1P,Past,Sg them to him/ give-1P,Past,Sg to him them

Another property of clitics is that they are deficient and cannot occur without the verb, while full NPs and pronominals can, as shown in (5a) and (5b, c), respectively:

(5) Q: Whom did (you/he/she) see?
Cl-Acc.Fem

b. Petar

c. Neja.
her

To the inventory of sentential clitics can be also added the reflexive clitic SE and pro as realizations of the subject clitic. Pro is the non-overt subject clitic, as in (6), while SE is found in passive constructions, as in (7), and as deficient arguments lacking person and number in ‘feel-like’ constructions, as in (8) (with the interpretation in (9)) (see Rivero 2001 for more discussion of the nature of SE in Romance and Slavic).

(6) Pro iznenada go.
surprised him-Cl.ACC
‘He/she surprised him.’

3 Full pronouns are also attested as salient dislocated elements.
(7) Tuk se raboti. 
*here Refl work.*
‘Here people work.’

(8) Na Ivan mu se stroi kashta. 
*to Ivan Cl-ACC Refl build house*

(9) Ivan feels like building a house (preferred reading: ‘feel-like’)

(10) A house is been built for the benefit of Ivan (passive reading)

The most natural claim about clitics appears to be that they are argument variables, which are never found in focus-accented final positions in the sentence. When the sentence contains only clitics, as in (4) or (6) above, there are no alternatives to the arguments, deriving in my view the ‘topicality’ effects discussed in the literature.

The controversy surrounding the argument/non-argument status of clitics has been handled differently for different languages, depending on advancements in the theory itself. For Bulgarian, Rudin (1997) and Franks & King (2000) adopt the view that clitics are functional agreement heads which may optionally associate with full DPs located in the VP.\(^4\) According to this view, the associated arguments appear in the usual thematic VP-internal positions (overt or null) ‘doubling’ the clitic. Unlike genuine object agreement markers, object clitics in Bulgarian are not, as this analysis predicts, always attested.\(^5\) It also remains mysterious why clitic doubling of arguments is not attested and the noun appears more often in the left periphery of the clause (with or without a clitic).

In a (2001) proposal, Bošković claims that pronominal clitics are actually generated in Specifier positions (being ambiguous X max/min elements)\(^6\) and the verb moves through empty heads, so that clitics left-adjoin to the verb in accordance with Kayne 1995. IO and DO clitics check phi-features against the same head but in distinct projections, and the verb and clitic cluster, consisting

---

\(^4\) See also Penchev 1993 and Dimitrova-Vulchanova & Hellan 1996 for the view that AgrIO and AgrO form a single constituent. Franks and Rudin (2004) revise this analysis and view clitics as K heads taking the noun as a complement. Clitics are claimed to be overt whenever their DP complement is silent, i.e., vacated (for various reasons, including TOPIC and pro as topic). They are silent when the DP is overt and in situ. This analysis would predict, however, that all topics are equal and require an overt clitic, i.e., it takes all topics to involve clitic reduplication. As will become clear from this paper, the semantic and distributional properties of contrastive topics are quite distinct from those of clitic-left-dislocated (inherent) topics.

\(^5\) Joseph (2001) and Kalluli (2001) provide convincing arguments against the agreement-marker hypothesis for other Balkan languages.

\(^6\) The idea that clitics are argument-like elements is not new. In Kayne 1975, object clitics are *pronominal arguments* generated in the canonical internal argument position of the verb.
of auxiliary and pronominal clitics, end up in the same head position (the highest projection in the inflectional domain).

In what follows, I adopt the essentials of this proposal without further discussion (and omitting technical details of the analysis that are not relevant to the claims made in this paper) and argue that clitics are argument variables while syntactically they are realized in Spec positions of verbal rather than agreement heads, as outlined in (11):

(11) Base-generation of clitics and pro prior to cluster formation under T/AgrP

\[
\text{VP pro/SE [V'} \text{ [VP mu [V'} \text{ [VP gi [V'} \text{ [VP V]]]]]]}
\]

4 Left-peripheral Elements in the Clause: Contrastive versus Dislocated Elements

In this section I distinguish CLLD from another construction in the left periphery of the clause — namely, contrastive topicalization — and present the general differences between the two: contrastive topicalization involves ordered pairs and feature valuing in T, while CLLD does not impose any pair ordering and is inherently salient/topical in nature. Clitics are absent in contrastive topicalization constructions, while they are obligatory in CLLDed structures.

4.1 Contrastive Topics

Contrastive topics (CT) participate in constructions which are similar to the English CT-F constructions discussed by Büring (2003), and can involve sets of ordered pairs, as in (12a), where the set of pairs is \{\text{Marija, hat}, \text{Milena, bag}\} and the CTs are the persons, while the F are the items purchased.

7 Another type of dislocation which I will omit from the discussion here is Hanging Topic Left Dislocation (HTLD), which is also possible. In HTLD, which roughly corresponds to Left Dislocation in English (see Cinque 1990 for relevant discussion), there is no syntactic connectedness between the dislocated element and the rest of the sentence and the dislocated element above the clause basically belongs to a different discourse unit. The dislocated element and the full pronoun inside the clause behave in this case more or less like two NPs in two different sentences.

(i) Brat ti, ama i toj e edin glupak.
   brother-Cl-Poss, EXLAM. and he i a fool.
   ‘Speaking of your brother, he is a fool.’

8 I exclude from the discussion cases where additional discourse-related contrast/emphasis and alternatives are evoked and consider here only contrast built into the grammar of the language.
Interestingly, in Bulgarian, the order subject-object can be reversed and constructions with D-linked objects as in (12b) are also attested. In this case the ordering in the pairs is reversed, the set becoming \{\text{hat, Marija}, \text{bag, Milena}\}, and the D-linked information is the objects, while the persons are the ‘new information’ answering the wh-question. As seen from the context questions, both orders in (12) involve CT on the higher element in the pair.

(12) a. Q: Who (from those people) bought what?
   What did Maria and Milena buy?
   A: \textbf{Marija-CT} kupi \textit{shapka-F}, a \textbf{Milena-CT} (kupi) \textit{chanta-F}
   \textit{Marija(Top) bought hat, while Milena (Top) (bought) bag}
   ‘Marija bought a hat and Milena — a bag.’

   b. Q: Who bought what?
   Who bought a hat and who bought a bag?
   A: \textbf{Shapka-CT} kupi \textit{Marija-F}, a \textbf{chanta-CT}(kupi) \textit{Milena-F}
   \textit{hat(Top) bought Marija, while bag (Top) (bought) Milena}
   ‘Marija bought a hat and Milena — a bag.’

As shown in (13), the contrastive (D-linked) element and the ‘focused’ element are obligatorily separated by the verb:

(13) a. *Shapka Marija kupi, a chanta Milena kupi.
   \textit{hat Mary bought, while bag Milena bought}

   b. *Marija shapka kupi, a Milena chanta kupi.
   \textit{Mary shapka bought, while Milena bag bought}

In each set in (12a, b), there is an F-marked constituent and a Contrastive Topic constituent characterized by a B-accent, similarly to English (see Jackendoff 1972 and Büring 2003).

One possibility not discussed by Büring but taken into account in this study is that the value for the F-marked constituent in CT constructions may be fixed for each ordered pair because it does not to involve alternatives in a set (see also the next section),\(^9\) while the value marking the contrastive topic varies, as in (14) (see also Cohen 2004 for a similar proposal for single ordered pairs):

(14) \{Mary bought hat, Ani bought hat, Milena bought bag, John bought bag\}.

This yields the correct meaning for universal statements with non-specific indefinites involving unique ordered pairs between two members, as in (15):

\(^9\) Büring (2003) adopts a view where in the pairs the focus value also alternates. See Cohen 2004 for more discussion.
(15) **Shokolad** obichat naj-mnogo dezata.

\[ \text{chocolate(Top) like the most children-the} \]

‘Children like chocolate the most.’ \{(chocolate, children)\}

In my view, the ordered pair in Bulgarian can also include non-argument members, as shown in (16) and (17) below. In this case, an argument which clearly does not raise for Case or for any other reason, as shown by the ungrammatical (16b) and (17b), raises to value the Contrastive feature (in the highest projection in the T domain, which I will assume here to be Spec,TP) and enters in a ‘pair relation’ with a place or time adjunct. This is a case of ‘topicalization’ of arguments on the condition that they participate in a pair where the higher element is contrastive, e.g. evokes other alternatives.

(16) a. Vali dâzdh (phi-features of ‘dazhd’ checked in situ)

\[ \text{rains/pours rain.} \]

‘It is raining.’

b. *Dâzhd vali. *CONTRASTIVE TOPIC

\[ \text{rain rains/pours} \]

‘It is raining.’

c. Dâdzh vali vav Winnipeg CT——F

\[ \text{rain rains/pours in Winnipeg} \]

‘It is raining in Winnipeg.’

(17) a. pro prochetoxme knigata. (phi-features of ‘kniga’ checked in situ)

\[ \text{(we) read book-the} \]

‘We read the book.’

b. *Knigata prochetoxme. *CONTRASTIVE TOPIC

\[ \text{book-the (we) read} \]

c. Knigata prochetoxme vchera CT——F

\[ \text{book (we) read yesterday.} \]

‘We read the book yesterday.’

4.2 Clitic Left Dislocation

Left-dislocated topics, on the other hand, are ‘inherent’ topics in the sense of Reinhart (1981), and pick out salient entities in discourse which are not necessarily contrastive and do not need to appear in a set — although nothing prevents them from being pragmatically contrasted with other entities. Most importantly, they do not involve ordered pairs involving the subject and the object, as shown below. Multiple CLLD elements do not answer wh-questions related to any of the dislocated elements in particular and their order can be reversed without any obvious consequences for the interpretation of the sentence.
(18) **‘LEFT-DISLOCATED’ TOPICS**

a. Ivan knigite vchera *(gi) vârna.
   Ivan (Top) books-the (Top) yesterday them-cl bring-Past

b. Knigite Ivan vchera *(gi) vârna.
   books-the (Top) Ivan (Top) yesterday them-cl bring-Past

‘Ivan returned the books yesterday.’

(19) **Context for (18 a, b):**

Possible context: What happened? (but with salient ‘Ivan’ and ‘books’ in mind)

It is true that Topic-Focus pairs are also possible with CLLD constructions\(^\text{10}\) when one of the elements is not dislocated, as also noted for Spanish by Arregi (2003). This is achieved when a focused element is found in the lower clausal domain, as in (20). In my view, however, these cases only superficially resemble their counterparts in the constructions with contrastive topics and involve ‘accidental’ pairings (knigite, Ivan)/(Ivan, knigite) of the salient topic and the focused element, since the ‘real’ pairing in this case is between the salient element and the whole predicate (gi varna Ivan) which is relevant to the interpretation of the topic (see section 7.1). Note also that the intonational properties of these constructions are quite distinct from those of the CT-F counterparts of these sentences without clitics.

(20) **LEFT DISLOCATED T (F is clause-internal)**

a. Knigite-T pro gi vârna Ivan-F.
   books-the (Top pro them-cl bringPast Ivan (Foc)
   (a risunkite –T gi vârna Emil-F.)
   while pictures-the Cl-Acc returned Emil.
   ‘Ivan returned the books (and Emil returned the pictures).’

b. Ivan-T pro vârna knigite-F.
   Ivan (Top) pro bringPast books-the (Foc)
   a Emil-T pro vârna risunkite-F.
   while Emil pro returned pictures-the
   ‘Ivan returned the books (and Emil returned the pictures).’

---

\(^{10}\) Recently it has been claimed that this construction is also attested in *wh*-questions, where an animate *wh*-word is resumed by a clitic (see Jaeger 2004 and Krapova and Cinque 2003 for examples and discussion):

(i) Kogo koj go narisuva?
   *whom who Cl-Acc painted?
   ‘Who painted whom?’ (Jaeger 2004)
In this section, I have shown that in Bulgarian, contrastive topics are semantically related to (sets of) ordered pairs and value a contrastive feature, while CLLD-ed topics are inherently topical in nature and are not ordered. They do not depend for their interpretation on a focused item found in the clause and are never focused themselves.

5 Contrastive Focus and Contrastive Topic: A Unified View

In what follows, I will propose that while the ‘pairing’ T-F in CLLD constructions is accidental, in CT-F (and Constrastive Focus-Topic) constructions, it is related to argument saturation and T-F relations in a ‘split focus’ fashion (see also Rooth 1985 and Jacobs 1984 for a discussion of discontinuous focus).

5.1 Contrastive Focus

My claim is that all contrastive constructions involve ordered pairs where the higher element is valuing a contrastive feature (cf. OCC in Chomsky 2001), while the element in the VP is a non-contrastive topic or focus.

Consider first the contrastive focus/non-contrastive (information focus) distinction, which is also discussed in Kiss’s (1998) study of Hungarian focus. In the first case, shown in (21), a set of alternative people (restricted or unrestricted) whom Ivan met is evoked, and one member, Marija, is exhaustively selected from all the alternatives. The answer is true if Mary was the only person Ivan met and nobody else. If there are other people whom Ivan met yesterday, the sentence is false.

(21) CONTRASTIVE (EXHAUSTIVE) FOCUS:

Marija-CF (*ja) posreshtnata vchera Ivan-T.
Mary- CF *CL-ACC met yesterday Ivan-T.
‘It was Mary Ivan met yesterday.’

The second case, shown in (22), exemplifies information focus, which is unnatural with a context question such as Is it Ani that Ivan met? The answer is also true if there are other people Ivan met in addition to Mary, clearly showing that Mary does not belong to a set of alternatives from which one is exhaustively chosen.

(22) INFORMATION (NON-EXHAUSTIVE) FOCUS:

Vchera (*ja) posreshtnata Ivan-T Marija-F.
yesterday *CL-ACC met Ivan-T Mary-F
‘Yesterday, Ivan met Mary.’
5.2 CF is the Reverse of CT

If we now consider CT, discussed in the previous section, then we can see that contrastive topics also involve alternatives found in a set on the higher element of the pair, with the difference that the wh-question is asking about the other member of the pair. A low-clause ‘Ivan’, on the other hand, as in (22) above, is not contrasted with other people and could be defined as the non-contrastive equivalent of information focus — non-contrastive topic. The alternative focus semantics of Rooth 1985 does not take into account non-contrastive focus (and topic), but Cohen (2004) defines a B-semantic value related to CTs and this is clearly a case where the topic value varies with respect to a fixed focus value, as shown in (23):

\[(23) \quad U [[[ x] B loves [y] F ]^B =
\{ John loves Mary, Fred loves Mary, …\} \quad (\text{Cohen 2004})
\]
\[
\text{(B = CT value)}
\]
\[
\text{CT: John; F: Mary}
\]

The proposal is, then, that this is the reverse of CF, where one alternative is linked to a non-contrastive topic value:

\[(24) \quad U [[[ x] CF loves [y] T ]^{CF} = \{ John loves Mary, Fred loves Mary, …\} \]
\[
\text{CF: John, T: Mary}
\]

On this view, foci and topics can be united under a ‘split’ chain hypothesis having a non-contrastive member in the vP-domain as the other member of the pair. This is shown in (25):

\[(25) \quad \text{a. CONTRASTIVE TOPIC \quad CT —— F (NON-CONTRASTIVE)}
\]
\[
\text{TP}
\]
\[
\text{CT}
\]
\[
\text{Value}
\]
\[
\text{VP}
\]
\[
\text{F}
\]
In all these cases feature valuing to T occurs (cf. OCC in Chomsky 2001). For example, in (26), the EPP/OCC feature of the object *eda kniga* is valued in Spec, T/Agr. This is shown in more detail in (27): the subject *Marija* remains in the vP and is the non-contrastive member of the pair.

(26) *Edna kniga prochete Marija.*

a book read ```Mary
'Mary read a book.'

Alternatively, both elements can remain in the vP/VP (and consequently keep their non-contrastive values). Their order can be also reversed by p-movement, as in (i), if the assignment of stress to the lowest element in the clause by the Nuclear Stress Rule (NSR) is not consistent with the value of the F-marked constituent, F₂ (see Zubizarreta 1998 for more discussion and Arnaudova 2001 for prosodic movement in Bulgarian).

(i) T_________F (Merge)

F₂_________T>T_________F₂ (P-movement)

A similar system is proposed for Korean in Choi 1996, where all [-prominent] elements correlate with lower positions in the sentence. Non-contrastive focus (‘completive’ in Choi’s study) also has the [+new] value, while [+prominent] elements, which appear in a higher position in the clause, can include either topical or contrastive focal elements, depending on the value of [new]. The proposed feature-based system is [-new/+prominent] for topics, [+new/+prominent] for contrastive foci, [-prominent/-new] for in-situ topics (tails) and [-prominent/+new] for in-situ foci (completive foci). See also Diesing’s (1992) ‘Tree-splitting hypothesis’, which partitions the sentence into a ‘nuclear scope’ and a ‘restrictor’.
Additional evidence for the existence of contrast/focus in a projection below C is found in wh-questions. I discuss these in the next section.

5.3 Contrastive (Focus) Feature in Wh-questions

In Bulgarian, all wh-words undergo obligatory movement to a functional projection (Spec,CP, as initially claimed by Rudin (1988)), while it would be sufficient, as in English, for just one of the wh-phrases to satisfy the requirements of the head C if the feature residing on the head has been already checked/valued. Examples of multiple wh-movement in main and subordinate clauses are shown in (28a) and (28b), respectively:

(28) a. Koj kakvo pravi?
   *who what does*?
   ‘Who does what?’

   b. Chudja se koj kakvo pravi.
   *wonder Refl who what does*
   ‘I wonder who does what.’

According to Bošković (1998), multiple wh-fronting in Bulgarian is an epiphenomenon consisting of focus movement for all wh-phrases and actual wh-movement for just one wh-phrase. In this analysis, Bulgarian resembles English in having wh-movement of a single phrase, all the other wh-phrases being fronted for independent reasons. Bošković proposes that the highest wh-phrase (the subject wh-phrase koj) moves to check the strong wh-feature of C, while the movement of other wh-phrases such as objects is actually Focus movement. He claims that the strong focus feature resides in the wh-word
instead of being on the target (which is FocP, in his view), while the \textit{wh}-feature checked by the first \textit{wh}-word resides in the functional projection \textit{C}. Every \textit{wh}-phrase with a strong focus feature must undergo focus movement, since the strong feature resides in the \textit{wh}-phrases, not in the target of the movement. Support for this account is found in the observation that there is a partial superiority requiring only the \textit{wh}-phrase that checks the strong +\textit{wh}-feature of \textit{C} (the subject \textit{wh}-phrase) to appear first, while the other \textit{wh}-phrases undergoing ‘pure focus movement’ can appear in any order, being insensitive to superiority, as shown in (29):

\begin{enumerate}
\item[(29)] a. Koj kogo kak e zelunal?
\textit{who whom how} \textit{Aux kiss-Past Part.}
\textquoteleft How did who kissed whom?\textquoteright

\item b. Koj kak kogo e zelunal?
\textit{who how whom} \textit{Aux kiss-Past Part.}
\textquoteleft How did who kissed whom?\textquoteright
\end{enumerate}

Given this, we can propose that the contrastive focus feature in \textit{wh}-questions and in declaratives has a common source and that contrastive elements target the same position. The most plausible candidate is the highest IP projection in the clause (T/AgrP or MoodP), which can be claimed to allow multiple specifiers with \textit{wh}-phrases.

\section{Semantic Type of the ‘Double’ and Syntactic Properties of CLLD}

In this section I provide evidence that CLLD topics are not permissible with non-specific indefinites, while CTs are. CLLDed elements are shown to be situated in a domain above the TP-level containing contrastive topic/focus.

CLLDed elements can be of various semantic types, as shown below: referential indefinites (30a), partitive indefinites (30b), and generic DPs, both indefinite (30c) and definite (30d).

\footnote{Clitic reduplication is not attested with any type of noun when both the clitic and noun ‘compete’ for the same focus/predication domain (in other words, there is no clitic doubling):

\begin{enumerate}
\item[(i)] ??Vidjaxa go choveka/nego/uchebnika/edin uchebnik/uchebnik
\textit{(they) saw CI-ACC man-the/him/textbook-the/a textbook/textbook.}
\end{enumerate}

If the verb or, more precisely the VP, is stressed and there is an intonational break after the verb, the sentence in (i) becomes acceptable. However, clitic right dislocations (CLRD) are not instances of clitic doubling, because the ‘double’ is actually not a thematic argument of the V (see also Warburton et al. 2004 for discussion). CLLD and CLRD are quite similar from the point of view of information packaging (but see Vallduví’s (1990) analysis of ‘link’}
(30) a. REFERENTIAL INDEFINITE:

Edin moj prijatel go vidjaxa da izliza ot xotela.
*one Poss friend Cl-ACC (they)saw DA (he) walks out of hotel-the
‘They saw of friend of mine leaving the hotel.’

b. PARTITIVE INDEFINITE:

Dve ot rozite gi podarixa na Marija.
*two of roses-the Cl-ACC give-3P,Past,Pl to Marija.
‘(They) gave two of the roses to Mary as a gift.’

c. INDEFINITE GENERIC:

Edna interesna kniga mozhe da ja chetesh zjal den.
*one interesting book can DA Cl-ACC (you)read whole day.
‘You can read an interesting book all day long.’

d. DEFINITE GENERIC:

Interesnata kniga ne mozhe lesno da ja ostavish nastrana
*interesting-the book NEGcan easily DA Cl-ACC leave aside
‘You cannot leave aside so easily an interesting book.’

There is, however, a restriction on CLLD elements: while indefinites are possible when specific, generic, or referential, as shown in (30), non-specific indefinites and bare plurals are never allowed in this position, as shown in (31). Non-specific indefinites are found only with contrastive topics or contrastive foci, as shown in (32).

(31) CLLD

a. *Paket go izgubixa uchenizite.

package Cl-ACC lost students-the
Intended: ‘The students lost a package.’

b. *Paketi gi izgubixa uchenizite.

package Cl-ACC lost students-the
Intended: ‘The students lost packages.’

and ‘tail’ for their differences). Often, CLRD is viewed as adjunction to the right of the same node (the so-called ‘mirror hypothesis’), but recently this view has been challenged for Catalan and Spanish (Villalba 2000, Cecchetto 1999) in line with Kayne’s system, where right adjunction is prohibited. Both Villalba (2000) and Cecchetto (1999) arrive at the claim that the ‘right’ dislocated element actually originates as a low specifier in the VP-periphery of the sentence. Warburton. et al. (2004) propose that in Greek right adjunction also occurs at the vP-level. I will leave this question open for further research.
a. Paketi izgubixe uchenizite.
   *packages lost students-the*
   ‘The students lost packages.’

b. Paket izgubixe uchenizite.
   *package lost students-the*
   ‘The students lost a package.’

In addition, I show that CLLDed structures have a number of other characteristics that distinguish them from CT/CF: lack of weak crossover (WCO) effects, no possibility of extraction, restricted island sensitivity, no scope ambiguity, and no reconstruction to thematic positions.

First, according to some native speaker judgements, CLLD in Bulgarian displays selective island sensitivity: it freely violates the wh-type of islands (33); while strong islands, such as complex NP islands (34) and adjunct constraints (35), can be violated only for some speakers.

(33) Knigata ne znae kakvo da ja pravi.
   *book-the Neg (he/she) knows what DA CL-ACC do*
   ‘He(she) does not know what to do with the book.

(34) ?Marija sreshtnax mazha kojto ja obicha.
   *Mary (I) met man-the who Cl-AC loves*
   ‘I met the man who loves Mary.’

(35) ?Vestnika zaspa dokato go cheteshe.
   *newspaper-the slept while Cl-ACC (he) read-Imp.Past*
   ‘He/she fell asleep while he/she was reading the newspaper.’

In (36a), the variable expression *vsjako dete* ‘each child’ can be coindexed with the pronoun *mu* ‘his’ without creating WCO effects, indicating that the former is not related to a lower copy/trace to the right of the subject *majka mu* ‘his mother’. In contrast, CT/CF constructions do give rise to WCO effects.

(36) a. [Vsjako dete] majka mu go obicha.
   *each child mother his Cl-ACC likes*
   ‘Each child is loved by his mother.’

b. *[Vsjako dete] obicha majka mu.
   *each child likes mother his*

Scopal ambiguity is not found with CLLD elements, a fact which can be explained in terms of an absence of reconstruction to thematic positions:
With CTs, a distributive marker *po* is used to disambiguate the sentence (otherwise the sentence is ambiguous between the two readings shown below):

(38) a. *Po* edna kniga prochete vsjako dete.  
   *DIST a book read each child  
   ‘Each child read (a different) book’

   *one book > each child  
   each child > a book

b. Edna kniga prochete vsjako dete  
   *a book read each child  
   ‘Each child read a book’

   one book > each child  
   each child > a book

CLLD cannot involve the quantificational expression *prekaleno mnogo* ‘too many’, while CTs and CFs can, as illustrated in (39) and (40), respectively:

(39) *Prekaleno mnogo knigi gi prochete Ivan.  
   too many books CL-ACC read-Past Ivan.  
   ‘Ivan read too many books.’

(40) Prekaleno mnogo knigi prochete Ivan.  
   too many books read-Past Ivan.  
   ‘Ivan read too many books.’

No extraction out of dislocated doubles is possible, while such extraction with CT/CF is quite common, as shown in (41) and (42), respectively:

(41) *Na Felini, go vidjax filma.  
   of Felini CL-ACC see-Past movie-the

(42) Na Felini vidjax filma.  
   of Felini see-Past movie-the  
   ‘I saw Felini’s movie.’

Now consider minimality effects with focused phrases. In the presence of a contrastively focused constituent in the left periphery of the clause, dislocated subjects need to be in a position above the focused constituent, as (43) shows:
(43) a. *TOZI PRINZIP Chomsky opisa.
   *this principle Chomsky described

   b. Chomsky, TOZI PRINZIP opisa.
   Chomsky *this principle described
   ‘It is this principle Chomsky has described.’

Similarly, dislocated objects cannot appear lower than contrastively focused subjects, as we see in (44):

(44) a. *IVAN Marija ja obvini
   Ivan Mary Cl-ACC accuse-Past

   b. Marija IVAN ja obvini.
   Mary Ivan Cl-ACC accuse-Past
   ‘It was Ivan who accused Mary.’

The minimality constraints and the syntactic evidence presented above suggest that the distribution of CLLD in the clause is as in (45):

(45) CLLD-Topic CT/CF *CLLD-Topic clitic/pro + verb VP

In this section I have provided evidence that dislocated elements are not base-generated in positions related to V and do not raise to Spec,T/Agr, which results in a number of syntactic differences with CT/CF structures. In addition, non-specific indefinites are banned from CLLD-ed structures while they are admitted in constructions involving contrastive topicalization/focalization. In the next section I explore the possibility that there are two types of predication in Mohawk and Bulgarian, overt DPs being either VP-internal (by incorporation) or base-generated in adjoined positions.

7 Two Types of Predication in Mohawk and Bulgarian

7.1 Mohawk (Baker 1996)

Any inflected verb in Mohawk is considered to be a complete and proper sentence as long as it occurs in an appropriate context. Verbs have rich inflection, displaying person, number, and gender features for both subject and object, as shown in (46). Intuitively, these inflections ‘count as pronouns, and provide the true subject and object of the verb’ (Baker 1996: 11). Such ‘pronouns’ may refer all by themselves — that is, even without the occurrence of any independent noun phrase within the sentence, as indicated in (46a). In this context, compare the Mohawk sentence in (46a) with the Bulgarian sentence in (46b), where the object argument is realized through a clitic.
SALIENCY AND CONTRAST IN COLLOQUIAL BULGARIAN

(46) a. Ra-NÚHWE’-s.
   Masc Sing.Subject—like—habitual marker
   ‘He likes it.’ (Baker 1996)

   b. Xaresva ja.
      like-3P,Pres,Sg Cl-Acc.Fem.
      ‘He likes her/it.’

Interestingly, overt NPs in Mohawk can precede or follow an initial ‘sentence’ like Ra-NÚHWE’s in a fashion strikingly similar to CLLD in Bulgarian. In each language, the sentence represents a nucleus, given in bold in (47)-(50), and overt nouns can surround it in any order to the left or to the right, resulting in the proposition ‘Sak likes her dress.’

(47) a. Sak ra-NÚHWE’-s ako-[a]tyá’ tawi.
   Sak MascSingSubject-like-habitual marker Fem.Sing.Poss. dress.

   b. Ivan, ja xaresva, rokljata j.
      Ivan, Cl-ACC like-1P,Pres,Sg dress-the Poss.

(48) a. Ra-NÚHWE’-s Sak ako-[a]tyá’ tawi.
   MascSingSubject-like-habitual marker Sak Fem.Sing.Poss. dress

   b. Xaresva ja, Ivan rokljata j.
      like-1P,Pres,Sg Ivan, Cl-ACC Ivan dress-the Poss.

(49) a. Sak ako-[a]tyá’ tawi ra-NÚHWE’-s.
   Fem.Sing.Poss. dress MascSingSubject-like-habitual marker

   b. Ivan rokljata j, ja xaresva.
      Ivan dress-the Poss Cl-Acc.Fem. like-1P,Pres,Sg

(50) a. ako-[a]tyá’ tawi Sak ra-NÚHWE’-s.
   Fem.Sing.Poss. dress Sak MascSingSubject-like-habitual marker

   b. Rokljata j Ivan, ja xaresva.
      dress-the Poss Ivan Cl-ACC like-1P,Pres,Sg

Baker (1996) argues that Mohawk and other Amerindian languages have two ways of realizing predication: (i) through the inflectional (agreement) morphemes expressing person, number and gender features which count as the subject and the object of the verb, as shown above; and (ii) through incorporation, where a root attaches to the verb but can still be viewed as the syntactic complement of the verb in examples such as ‘fish-prepare’. According to Baker, agreement morphemes and incorporated noun roots are part of the same system, since they express theta-roles and are head elements. This is stated in the Polysynthesis parameter, understood as a condition on theta-role assignment:
A phrase X is visible for theta-role assignment from a head Y only if it is coindexed with a morpheme in the word containing Y via:

- an agreement relationship (e.g. between a morpheme and an NP);
- a movement relationship (incorporation).

The first case, explored by Baker, is represented by so-called agreement morphemes. They are argued to be syntactic heads, as shown in (52), and express theta-roles. Full NPs, when they appear, have an adjunct or modifier status.

(52) AGREEMENT MORPHEMES IN MOHAWK (Baker 1996)

```
IP
    NP1
       Sam
    I'
       I
          VP
          V
             NP2
                her dress
```

The second possibility, incorporation, is exemplified in (53) for a sentence such as *I am buying a bed* (lit. bed-buy). Baker proposes an English-like complementation structure with the head of the noun undergoing syntactic movement, and shows that the roots, which incorporate, are arguments in their own right. That is, the theta-criterion is met syntactically by an incorporating noun and the dislocated NP may be phonetically empty.

---

13 This parameter combines the pronominal argument hypothesis initially proposed by Jelinek (1984), which basically states that the Theta-criterion is met morphologically and inflectional (agreement) morphemes expressing person, number and gender features count as the subject and the object of the verb and Baker’s own incorporation hypothesis, according to which an incorporated root attached to the verb is actually the syntactic complement of the verb.
In (53), the object, which is incorporated, corresponds to a null, trace-headed NP in object position. There is no subject incorporation, however, and in this case, the subject morpheme corresponds to a null NP in subject position.

### 7.2 Proposal for Bulgarian

From the discussion in the previous section, we can conclude that there are two ways of achieving saturation across languages: arguments are realized either as full-fledged nouns (although in incorporating languages, only the object can appear as a full noun in the lower predication) or as inflectional markers/argument variables. In what follows, I propose that in Bulgarian arguments can be base-generated at two different levels: a lower level within the vP/VP and a higher level above IP (TP), which also participates in the saturation of the predicate. The higher level contains full NPs or pronouns, which saturate the predication containing clitic variables. Topic-focus chains containing contrastive elements are found only in the lower domain and saturate the predication ‘directly’.

In previous analyses of CLLDs in different languages (see Cinque 1990; Iatridou 1991, among others), the left-hand noun is understood as the subject of predication, and takes a predicate containing a variable, the clitic, an open position that permits a constituent to behave as a predicate:

(54) \[ [NP \text{DP} [\text{IP cl} \ldots]] \]

Consider now (55), where the NP Ivan can be either a dislocated object, as in (a), or a subject, as in (b):
(55) a. Ivan go vidjaxa.
   *Ivan CL-ACC (they) saw*
   ‘As of Ivan, they SAW him.’

   For some x (x = Ivan) they saw x.

b. Ivan dojde.
   *Ivan come-3P,Past,Sg.*
   ‘Ivan came.’

   In (55), the referent is picked up again by the description, similar to what we find with a so-called E-type pronoun and its antecedent (see Evans 1980):

(56) there is an x (x=Ivan)

   the x (such that x = Ivan) came/was seen etc.

   While CTs and CFs are linked to internal restrictor domains, external restrictor domains define a separate background existential presupposition related to discourse and identifying an entity (inherent ‘topic’ or event). This external domain has been equated semantically with ‘subject of predication’ (Reinhart 1981), ‘higher predication domain’, and ‘argument externalization’ (Zubizarreta 2000; Arnaudova 2001, 2003), and has described in syntactic terms as realized by adjuncts (see, e.g., Warburton et al. 2004) or by elements occurring in the specifier positions of topic operators (Zubizarreta 2000).

   I propose that CLLD and contrastive structures in Bulgarian, as discussed in this paper, exemplify two different types of argument saturation. In the case of CLLD, saturation is achieved through clitic variables but is incomplete (compare restricted saturation types discussed in Chung and Ladusaw, in press); consequently, a ‘double’ can properly saturate the predicate.

(57) [_{CLLD} Petâr na Marija edna kniga [_{TP} pro [_{TP'} j dade]]].
   *Peter to Marija a book pro Cl-DAT Cl-ACC gave.*
   Lit. ‘Peter to Mary a book, he gave it to her.’

(58) \(\lambda x \lambda y \lambda z (x \text{ gives } y \text{ to } z)\) <e, t>
   \(\exists x, x = \text{Petâr }\ldots\) <e>
   \(\exists y, y = \text{book }\ldots\) <e>
   \(\exists z, z = \text{to Mary }\ldots\) <e>

   This explains why non-specific indefinites of semantic type <e, t> cannot saturate the predicate:

(59) *Non-specific indefinites <e, t>*

   As a result, the dislocated element is an argument but is felt to be ‘removed’ from the domain of the predication, providing an independent description of the referent. The presence or absence of the dislocated element
does not alter the focus-topic structure of the lower predication domain, which is on the event or on an internal argument inside it.

8 Summary

In this paper I have shown that the left periphery in Bulgarian contains CLLDed elements and contrastive topic/focus elements found in a lower domain. Clitic left dislocation constructions have a number of syntactic and interpretative properties which make them closely related to a similar construction found in polysynthetic languages (Baker 1996) and represent a case where a ‘double’ properly saturates a function. In addition, I have offered a unified approach to the relation between contrastive focus and topic in the clause. The general claim is that a contrastive feature semantically interpreted as a set (restricted or unrestricted) containing alternatives is projected in the TP domain. Contrastive Topic-Focus and Contrastive Focus-Topic orderings are analysed as (sets of) ordered pairs (in the sense of Büring 2003 and Cohen 2004). The realization of focus in this case is related to direct argument saturation in the vP and TP-cycle in a ‘split focus’ fashion.

References


Franks, Steven and Catherine Rudin. 2004. Bulgarian Clitics as K^0 heads. Paper presented at Formal Approaches to Slavic Linguistics 13, University of South Carolina.


