

Iliyana KRAPOVA (Venice)

WORD ORDER AND DISCOURSE: THE CASE OF THE BALKAN LANGUAGES

Abstract: The paper investigates word order in three Balkan languages (Bulgarian, Greek and Romanian). It is argued that Balkan languages are discourse prominent and variations with respect to discourse notions like topic or focus are examined from a comparative perspective and in view of their hierarchical ordering in the left periphery of the sentence, which corresponds to that portion of the sentence which hosts topicalized and focalized constituents.

Keywords: topicalization; dislocation; clitics; focus; word order; Balkan languages

1. Introduction

Balkan languages can be said to belong to the so-called discourse-prominent languages, i.e., languages whose surface structure encodes through special syntactic means, rather than just prosodically, discourse(-semantic) functions such as Topic (discourse given or old information) and Focus (discourse new or emphatically represented information). In this contribution, I will show that at least in the three Balkan languages under study (Romanian, Bulgarian and Modern Greek, henceforth Greek), the two main types of discourse structures - topicalization and focalization – share a whole array of common syntactic properties and that their word order, at least in the preverbal field, is to a large extent shaped by information structure requirements.

We will use the terms ‘Topic’ and ‘Focus’, and respectively, *topicalization* and *focalization*, since they provide a useful conceptual basis for cross-linguistic generalizations regarding the relation between information structure and syntax. The terms themselves do not coincide with the traditional distinctions ‘Theme’ and ‘Rheme’, although their essence captures the traditional Prague school intuition that each sentence can be divided into a discourse-familiar or discourse-given part (*theme, osnova, základ*) and a discourse-new part (*rheme, jádro* ‘nucleus’, cf. e.g., Цыхун / Сухун 1962; Иванчев / Ivanchev 1978).

Given the correlations between types of phrases in the preverbal field, as well as their relative order, the purpose of this contribution is to show that the so-called ‘Left Periphery’ (cf. Rizzi 1997 and seq.) of the Balkan sentence is organized in a very similar way. Minimal variation between discourse structures is related to independent language internal differences, such as Case distinctions, the position of the clitic pronouns, use of special “prepositions” for object reduplication (such as *pe* in Romanian), etc.

Topic- and Focus-related notions can also be marked in sentence final positions, which alongside the sentence initial position, is endowed with discourse features (called ‘strong positions’ in Цыхун / Cyxun 1962: 268). The sentence final position is typically associated with one type of Focus, namely New Information Focus (cf. Kiss 1998) or *rheme* in the classical sense of the term. Following it, one can also find topicalized elements (direct and indirect objects) typically clitic doubled (‘anticipatio’, cf. e.g., Лопашов / Lopashov 1978)¹ or displaced to the right periphery of the clause (often labeled ‘Clitic Right Dislocation’ (Samek-Lodovici 2015). Cf. one example from Bulgarian: *Ostavi ya onaya Marian - poznavam ya dobre* ‘Leave her alone, that Marian – I know her well’ where the verb in the first sentence carries main stress, and the dislocated object doubled by an agreeing clitic is deaccented precisely to indicate discourse-givenness. Comparable cases in other languages (Spanish, Italian...) have been argued to involve an extra-sentential position for the right-dislocated phrase (Ott & De Vries 2014). Such structures are still in need of a better understanding given the pervasive use of such constructions in all Balkan languages (for a discussion on Greek and some references, see Philippaki-Warbuton et al. 2004).

In this paper, I will only concentrate on the Left Periphery of the clause, following illuminating work by Rizzi (1997), which has not yet been applied in a systematic way to the study of word order and information structure in Balkan comparative syntax.

2. The position of Topic and Focus in the left periphery of the Balkan languages

A sentence-initial Topic or Focus position is typical for all three Balkan languages under study. Here are some illustrative examples:

- (1) a. Ivan/nego ne sam go vizhdala otdavna. (Bul)
 ‘(As for) Ivan/him, I haven’t seen him for a long time’
 b. **Samo Ivan/nego** shte pokanya na sreshtata.²
 ‘It is only Ivan/him I will invite’
- (2) a. Tin Eleni/afti dhen tin idha (Gr)
 ‘(As for) Eleni/her, I haven’t seen her’
 b. **To Jani/afton** thelo na kalesis
 ‘It is Jannis/him that I want you to invite’

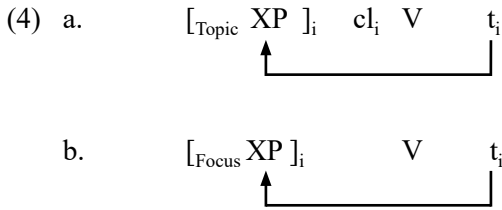
- (3) a. Pe Ion/el, l-am invitat³ (Rom)
 ‘(As for) Ion/him, I invited him’
 b. Numai pe Ion/el l’am invitat (Alboiu 2002: 260, ex. (22))
 ‘It is only Ion/him I invited’

The examples in (1a, 2a, 3a) involve topicalized direct objects (a proper name and a tonic pronoun); the examples in (1b), (2b), (3b) involve the same types of direct objects appearing as focalized phrases.

Reduplication by a pronominal clitic (also referred to as ‘reprise’) is the classical mark of object topicalization in all the Balkan languages (Лопашов / Lopashov 1978; Цыхун / Cyxun 1981; Асенова / Asenova 2002).⁴ From a formal syntactic point of view, Topic structures are seen as involving **dislocation** of an (direct or indirect) object to the preverbal position. From an information structure point of view, the word order corresponding to (1a), (2a), (3a), is referred to as *objective word order* (‘prav slovoed’, cf. Иванчев / Ivanchev 1978), since sentence initial Topics are linked to the preceding discourse and thus serve as a starting point (‘terme de départ’, in Guéntcheva’s 1994 terminology) for the actual predication. The Topic can also be viewed as the logical (notional) subject of the predication, i.e., what the predication is about. The rest of the sentence belongs to what is generally called ‘Comment’, i.e., the predication itself (cf. Vallduvì 1992). Since the clitic is obligatory, this type of dislocation has also been termed **Clitic Left Dislocation** (CLLD) (a term introduced by Cinque 1977 for similar constructions in Romance involving a short coreferential pronoun).

Focalized phrases, on the other hand, enter another type of information structure articulation: the Focus – Presupposition articulation, well-known since Chomsky (1973). Sentence initial Focus (also referred to as ‘Contrastive Focus’ or ‘Identificational Focus’, cf. Kiss 1998) is a specific type of focus.⁵ Pragmatically, it expresses the speaker’s intention to resolve a potential misunderstanding or doubt on the part of his interlocutor, or to correct some (part of a previous) statement. Therefore, Contrastive Focus is typically associated with some contextually determined set of alternatives for which the predicate holds potentially, by pointing out the unique member (or subset) of that set for which the predicate actually holds (Zubizarreta 1998: 6). Syntactically, this strategy makes use of the *subjective word order* (‘obraten slovoed’, cf. Иванчев / Ivanchev 1978): the information is presented as the most relevant part of the utterance and is typically pronounced with (strong) emphasis, i.e., it carries emphatic stress (‘logichesko udarenie’, cf. Попов/Попов 1961; Цыхун / Cyxun 1962: 287). This type of Focus conveys new information only indirectly: by emphasizing the information the speaker typically brings forward a (potentially) novel quality or property of what is being talked about, i.e., of the discourse theme.

Given the examples in (1b) - (3b), Focus can also be said to involve **dislocation**, but without an accompanying clitic pronoun. The dislocation of a Topic or a Focus to a preverbal position can be schematically represented as in (4), a & b respectively.



The abstract representations in (4) indicate that topicalization and focalization involve the same type of structure but differ in the presence vs. absence of a clitic (though not in Romanian, where focused constituents too can be doubled with a clitic when referring to [+human] entities, cf. (3b), section 3. below). In both (4a) and (4b), the object starts out from an object position, as a verbal argument, and dislocates to the preverbal position, leaving a trace (t_i) in its original position. Only in (4a), the agreeing clitic mediates the syntactic relation between the preposed object and its trace, ensuring co-referentiality (Guéntcheva 1994: 119).⁶

In case the dislocated element is an adverbial phrase, preposing to a sentence initial position never involves clitic doubling, given the absence of adverbial clitics in the Balkan languages, so the differentiation between a topic or a focus structure is achieved only prosodically (low stress, flat intonation vs. emphatic stress, high pitch or marked intonation). In the absence of such clues, potential discourse ambiguities are resolved by context (cf. Joseph & Philippaki-Warbuton 1987: 99-102 for Greek; Rudin 1991 for Bulgarian). See the following examples from Greek and Bulgarian:

- (5) a. [Sto xorio] tis pijeni poli sixna
lit. 'To her village she goes often'. (Gr)
[Me sevazmo] prepei na milate sto patera sas
'With respect you have to talk to your father'.
[Stu Jani] na pame apopse.⁷
'To Jannis' place let's go tonight'
- b. [Predi nyakolko dni] beshe hodila iz selo peperuda (Bul)
'Several days ago a butterfly was flying across the village'
[Vav vsyako hudozhestveno proizvedenie] tryabva da ima dvizhenie
'In every fiction, there must be some kind of action' (AG/AG 1994: 176)

2.1. Two types of topicalization structures in the left periphery

Two types of Topic structures can be distinguished in the Balkan languages under study: alongside (1a)-(3a), which present a clear case of CLLD structure, there are cases ((6)) where we find an additional emphatic tonic pronoun matching the phi-features and the morphological case of the associated clitic (as the identical indices illustrate). Resumption with an emphatic pronoun is optional but the possibility of

inserting a second associate for the topic constituent raises the question of how such constructions differ from ordinary CLLD:

- (6) a. (Cît despre) Ion_i, nu l_i-am văzut (**pe el_i**) de anul trecut.
 as for Ion pe he.acc not he.acc-have.pr.1s see.ppl of year last
 ‘(As for) John, I haven’t seen him since last year’
- b. (Kolkoto do) Ivan_i, vchera go_i vidyax (**nego_i**) da pie bira
 v edin bar. (Bul)
 ‘As for Ivan, I saw him yesterday drinking beer in a bar’
- c. (Oson afora tin) Maria_i, dhen tin_i anteho (**afti_i**) allo. (Gr)
 ‘As for Maria, I can’t stand her anymore’
 (Alexiadou 1997, 69, ex. (46a))

A coreferential tonic pronoun can never co-occur with the topics in (1) – neither as the sole resumptive element nor as an additional associate of the clitic. This fact, illustrated in (7), has been noted independently for each Balkan language: see Dobrovie-Sorin (1994: 218f) for Romanian; Rudin (1986: 32ff); Arnaudova 2001, 2010:165f; Krapova & Cinque 2008 for Bulgarian;⁸ Alexiadou (1997: 70); Anagnostopoulou (1994, 1997) for Greek. Rudin (1986: 34) also notes (for Bulgarian) that a full pronoun in (6b) can be missing only when it can be missing in a main clause as well. In other words, what follows the topic phrase *Ivan* in (6b) is a complete sentence with no constituent lacking that would have to be present in a normal sentence. This strongly suggests that the topic constituent of (6) occupies a clause-external position,⁹ while the same topic constituent occupies a clause-internal position in (7):

- (7) a. Pe Ion_i l_i-am întâlnit (* pe el_i) anul trecut. (Rom – Sorin 1994:218)
 Pe Ion him.CL have met (*him) last year’
- b. Ivan_i go_i sreshtnax (*nego_i) minalata godina. (Bul)
 Ivan him.CL met.1SG him last year
 ‘John, I met last year’
- c. Ti Maria den tin anteho (*afti) allo (Gr – Alexiadou 1997:70, ex. 46b)
 the Maria not her.CL stand.1SG her anymore

The Romance construction equivalent to that of (6) have been first studied by Cinque (1977, 1990) and has been labeled **Hanging Topic Left Dislocation (HTLD)** – a term which was meant to distinguish it from the CLLD construction analogous to the one in (7) (see also Kiss 1995). The chief property of the Romance HTLD is that the topicalized constituent, even though definite or given, entertains a rather loose relation with the following Comment. i.e., the HT creates only a general context for

the Comment, so it is often employed with the function of shifting an old topic to a new one (Topic shift). According to Krifka (2001), such usage constitutes “a speech act by itself, an initiating speech act that requires a subsequent speech act like an assertion, a question, command or curse about the entity that was selected” (p. 25). See also Guéntcheva’s 1994 term *extraposition Topic, a segmented phrase* (in the sense of Bally 1932/1965).

Alongside the type of resumption (i.e., the availability of a tonic pronoun inside the clause), another major property of the Balkan HTLD is the absence of a syntactic relation with the following clause as also signaled by the *as-for* expressions (*shto se otnasya do/kolkoto to, cît despre, oson ja/oson afora*), whose purpose of to close off the HT from its Comment.¹⁰

2.2. Properties of Hanging Topics

The special pragmatic properties of HTs are marked prosodically: often, there is a sharp intonational break between the left dislocated phrase and the rest of the sentence,¹¹ especially when the topic is introduced by an *as for* expression (‘thématisateur’ in Feuillet’s 1990 terminology):

- (8) A sarceto # byas go kasa **nego** kleto (Bul - Guentchéva 1994: 157)
 and the-heart anger it-CL.ACC tears it poor
 ‘And the poor heart, anger tears it apart’

The presence of a bigger intonational break or of a pause between the extra-sentential topicalized element and the following sentence has also been reported for Greek (9a) (Alexiadou 1997: 69, Anagnostopoulou 1997) and for Romanian (Dobrovie-Sorin 1990). Such a break or pause is missing in the respective CLLD construction (9b):

- (9) a. Ton Petro_i # ton_i nostalgo ton_i gliko mu¹² /afton_i poli zontas makria
 (Gr – HTLD)
 the Peter him_{cl’acc} miss the sweet my/him_{acc} much living far away
 ‘Peter, I miss my sweetie a lot living far away’
 b. Ton Petro_i ton_i nostalgo poli (Gr – CLLD)
 the Peter him_{cl acc} miss much
 ‘I miss Peter a lot’ (Anagnostopoulou 1997: 153)

The second and perhaps the most important diagnostic has to do with *Case connectivity*, which are often absent in HTLD constructions (Cinque 1990 for Romance; Bertolussi 2017 for Latin). In Greek, for example, the dislocated phrase often appears in the nominative case (*Nominativus pendens*)¹³ resulting in a mismatch with respect to the (accusative) case features of the resumptive element.

Co-referentiality in this case is signaled only by the shared agreement features. Cf (19) from Greek:

- (10) I Maria # tin ematha kala tosa xronia, ksero pos na tis miliso.
 the Maria_{NOM} her_{cl.acc} (I) knew well so-many years, (I) know how to
her_{cl.acc} talk
 ‘Maria, I have figured her out after so many years, I know how to talk to her’
 (Anagnostopoulou 1997: 154)

In CLLD, on the other hand, as seen in (9b) above, case-agreement is obligatory between the topic and the doubling clitic, so the dislocated XP acts as a real double of the clitic since prior to movement (see the structural representation in 4a) XP occupies the argumental position and is co-indexed with the clitic. The difference in case connectivity strongly suggests that HTs are base generated in their surface position rather than moved there.

Case connectivity effects are harder to see in Romanian and Bulgarian, since a Nominative vs. Accusative distinction is only visible in the pronominal system of these languages. Other diagnostics can also be taken into consideration. Dobrovie-Sorin (1990) suggests that in Romanian topicalized phrases introduced by *pe* can only enter the CLLD construction because this preposition (similarly to the differential object marker *a* in Spanish – but see note 4.) is licensed sentence-internally. This proposal receives support from the incompatibility between left dislocated objects introduced by *pe* and an emphatic pronoun typical for the hanging topic construction. See (11):

- (11) *Pe Maria nu vrea s-o mai văd **pe ea** cât trăiesc. (Rom)
 ‘Maria, I don’t want to see her as long as I live’
 (Dobrovie-Sorin 1994, 219, fn, 47, ex. (ii))

Under the assumption that *pe Maria* is base generated within the clause and can undergo movement into the left periphery whenever it is part of the CLLD construction (7a), the ungrammaticality of (11) suggests that whatever material appears in the pre-posed position must be base-generated there. In other words, Left Dislocation (in Dobrovie-Sorin’s terms) corresponding to what we here call HTLD, does not rely on movement.

For Bulgarian as well there are clues as to the movement of the preposed topic vs. its base-generation in an extra-sentential position. One such clue is the possibility of omitting the preposition/case marker *na* ‘to’, a phenomenon also known under the name of *na*-drop (cf. Vakareliyska 1994; Krapova & Cinque 2005). In terms of the present discussion, *na*-drop seems available if the respective indirect object is a hanging topic but not if it is a clitic left dislocated topic. This difference is illustrated in (12). In (12a), where an intonational pause is highly preferred (though

not obligatory), the *na*-less HT corresponds to a simple NP, base-generated directly in the preposed position; in (2b) on the other hand, the CLLD topic must be fully represented clause-internally before it can undergo movement, so *na*-drop becomes unavailable in this context. As expected, an emphatic pronoun is only possible in (12a) irrespective of its exact position:

- (12) a. Mariya# [az (na neya) tolkova chesto sam ì pomagal, (Bul–HTLD)
 Maria, I (to her) so often am her_{cl.dat} helped
 che tryabva da mi e blagodarna cyal zhiivot].¹⁴
 that should to me_{cl.dat} is thankful all life
 ‘Maria, I have helped her so often that she should be thankful to me for
 the rest of her life’
- b. [Na Mariya az (*na neya) tolkova chesto sam ì pomagal (*na neya)
 (Bul – CLLD)
 to Maria I (to her) so ofetn am her_{cl.dat} helped (*to her)
 che tryabva da mi e blagodarna cyal zhiivot].
 that should to me is thankful all life
 Lit. ‘To Maria, I have helped her so often that she should be thankful to
 me for the rest of her life.’

Given the contrast in (12), we can generalize that in the CLLD construction, the topic can be of any maximal projection, including a prepositional phrase, corresponding to the indirect object, while in the HTLD construction the dislocated phrase must be an NP. Since in Romanian there are no prepositional indirect objects, while Greek prepositional indirect objects cannot be clitic resumed (cf. *Sto Jani tha dhosi i Maria ta lefta avrio* ‘To John Mary will give the money tomorrow’ – Joseph & Philippaki-Warburton 1987: 99), then it must be the case that topicalized and Case marked indirect objects in these two languages participate only in the CLLD construction, observing Case connectivity.

A third diagnostic has to do with root-non-root character of the dislocated constructions. Anagnostopoulou (1994, 1997) and Alexiadou (1997) show that Nominative topics in Greek are never licensed in embedded clauses, so for example (10) above, if embedded, becomes ungrammatical, cf. (13).

- (13) Ipe oti *i Maria/ tin Maria tin emathe kala tosa xronia.
 (he/she) said that the Maria_{NOM}/the Maria_{ACC} her_{cl.acc} (he/she) knew well
 so many years
 ‘He/she said that he/she had figured out Maria after so many years’
 (Anagnostopoulou 1997: 154, ex. (6))

Since there is a parallel restriction in Romance (Cinque 1977), we can view the obligatory root nature of the clause as a third general diagnostics for detecting the presence of a HTLD construction. Consequently, whenever a configuration of the type $XP_i \dots cl_i$ is found in an embedded clause, it should automatically qualify as CLLD structure provided it can be shown to result from a movement derivation. In Bulgarian, only root contexts correlate with the availability of *na*-drop on the topicalized constituent. See the examples in (14).

- (14) a. Ivan kaza, che *(na) Mariya, toy izobshto ne i_i e pomagal.
 Ivan said that (to) Maria he at all not her_{cl.dat} is helped
 ‘Ivan said that he hadn’t helped Maria at all’
- b. Ivan # ne pomnya dali na nego (mu) dadohme neshto?
 Ivan, not (I)remember if to him (him_{cl.dat} (we)gave anything
 (Rudin 1986: 36, (45c))
 ‘Ivan, I don’t remember if we gave him anything’
- c. *Ne pomnya Ivan # dali na nego (mu) dadohme neshto
 not (I)remember Ivan if to him (him_{cl.dat} (we)gave anything
 (Rudin 1986: 36, (45d))

2.3. Pronominal Topics in Bulgarian

Krapova & Cinque (2008: 259ff) point out that Bulgarian colloquial speech makes an abundant use of Nominative-marked pronouns as hanging topics in root clauses (15) and that a hanging topic may not only coexist with a clitic left dislocated topic but if they do, they must obey a certain order as visible from the pronominal morphology and case connectivity effects (16). In (15) the first topic (*toy*) is nominative-marked (*Nominativus pendens*), while in (16) the same topic is followed by a dative- or accusative-marked tonic pronoun which agrees in case with the clause-internal doubling clitic:¹⁵

- (15) Toy ne mogat da go prikrepayat kam nikogo.
 he not (they)can to him_{cl.acc} (they)attach to anybody
 ‘As for him, they cannot attach him to anyone’
- (16) a. Toy # na nego tova nikoga ne mu se e sluchvalo.¹⁶
 he to him this never not me_{cl.dat} refl is happened
 ‘As for him, this has never happened to him’
- b. Toy # nego oshte go e yad, che navremeto ne se zapisal da sledva
 he him still him_{cl.acc} is anger that at the-time not enrolled to study
 ‘As for him, he still regrets that he didn’t enroll in University at the time’

In light of this brief discussion, the ambivalence of hanging topics should be compared to the *as-for*-topics which are never integrated in the clause and appear to always be extra-sentential. *As-for*-topics can co-occur with pronominal topics as well as with clitic left dislocated topics (in the order *kolkoto do NP* > *pronominal topics* > *clitic left dislocated topics* – see (21). Both CLLD and HTLD phrases count as occupants of the first position with respect to Tobler-Mussafia effects (proclisis vs. enclisis), while *kolkoto do*-phrases do not – see (20).

- (20) *Kolkoto do mene# me pokanixa na sreshtata oshte vchera.
 as for me mea_{clcc.} (they)invited to the-meeting already yesterday
 ‘As for me, they invited me to the meeting already yesterday’
- (21) Kolkoto do Ivan, [_{HT} toy [_{CLLD} nego [_{Subject} nikoy ne go haresva
 as for Ivan, he him nobody not him_{cl.acc} likes
 ‘As for Ivan, nobody likes him’

3. Linear orders between multiple topics

The general property of HTs (namely, that they occupy an absolute sentence initial position) is supplemented by a uniqueness requirement: there can only be a single HT per sentence. CLLD Topics, on the other hand, are exempt from the uniqueness requirement so in principle more than one dislocated phrase can appear clause-initially. Cinque (1977) also points out for Romance that there is practically no theoretical limit to the number of phrases that may undergo CLLD. The data collected by Alboiu (2000) for Romanian, by Anagnostopoulou (1994, 1997) for Greek, and by Arnaudova (2010); Krapova (2002); Джонова / Dzhonova (2004) for Bulgarian confirm this generalization, although there are pragmatic constraints on the order of multiple topics. For example, according to Alboiu (2000: 270), in Romanian, the highest Topic has maximum relevance for the discourse context, but otherwise all combinations are possible. In the examples below, Topics are given in brackets, so that their free ordering can be made more evident. Example (22) is adapted after Anagnostopoulou (1997: 160, ex. (21a)):

- (22) a. [_TTa vivlia] [_Ttis Marias] tis ta edhose o Jannis ‘ (Gr)
 the books to Maria her_{cl.gen} them_{cl.acc} gave the Jannis’
 [_TTis Marias] [_Tta vivlia] tis ta edhose o Jannis.
 To-the Maria the books her_{cl.gen} them_{cl.acc} gave the Jannis
- b. [_TMioarei] [_Tinelul] la nuntă i l-a dat Anghel (Rom)
 Mioara_{dat} ring-the at wedding her_{cl.dat} it_{cl.acc} -has given Anghel
 [_TInelul] [_TMioarei] la nuntă i l-a dat Anghel
 ring-the Mioara_{dat} at wedding her_{cl.dat} it_{cl.acc} -has given Anghel
 (Alboiu 2002: 270, ex. (32))

- c. [_TNa Mariya] [_Tpismoto] i go dadox az (Bul)
 to Mariya letter-the her_{cl.dat} it_{cl.acc.} gave I
 [_TPismoto] [_Tna Mariya] i go dadox az.
 letter-the to Mariya her_{cl.dat} it_{cl.acc} gave I

For Bulgarian in previous work (Krapova 2021) I have shown that the absolute first instantiation of Topic, i.e., the leftmost one, fits into what Bianchi & Frascarelli (2010) label A(boutness) Topic (also used for topic shift alongside a hanging topic): this type of topic signals that the speaker wants to introduce the topic constituent as salient for the purposes of the conversation or to reintroduce it into discourse as a new topic to be commented on in the future conversation. The second topic in the sequence can be said to perform what Bianchi & Frascarelli (2010) label a G(iven)-Topic and is typically an instance of a familiar or given topic in the sense of Schwarzschild (1999) whereby the topicalized constituent is not highlighted but simply refers back to an antecedent known from previous discourse. In this case, the topic resumes background information shared by the discourse participants.

3.1. Movement of the clitic left dislocated topic

Recall that in section 1, we postulated that the Left Periphery of the sentence contains a left dislocated topic construction position which is targeted by clitic resumed material counting as a Topic. We also argued that the hanging topics (leaving aside the case of pronominals) are not derived by movement from a base-generated argumental position. This view seems to be supported by all authors working on the distribution of HTs in the Balkan languages (Rudin 1986; Dobrovie-Sorin 1990; Anagnostopoulou 1997; Alexiadou 1997). Some arguments to this effect are presented below. As far as CLLD is concerned, Case Connectivity already indicates a movement derivation: the matching clitic functions as an anaphoric element connecting the original (base) position of the argument to its surface dislocated position.

3.1.1. ‘Unboundedness’

A first piece of evidence that the CLLD construction is derived through movement comes from the fact that it is not limited to monoclausal domains (Anagnostopoulou 1997): the dislocated phrase can appear outside of the embedded clause to which it belongs. Hence the term ‘unboundedness’. (23) provides examples from Greek and Bulgarian showing that the embedded Topics can also be dislocated into the domain of the matrix clause.

- (23) a. Tin Eleni_i su ipa xthes oti t_i tin idha t_i. (Gr)
 ‘Eleni, I told you yesterday that I saw her’
 b. Prestapnika_i mislya, che t_i sa go hvanali t_i. (Bul)
 ‘The criminal, I think they got him’

Such observations point that a movement operation takes place – the Topic starts out from the complement clause and dislocates to a position in the Left Periphery of the embedded clause, after which it moves into the Left Periphery of the matrix clause. This is indicated by the identical indices on the traces left at the positions which the topic passes through on its way to its dislocated surface position. One such position is immediately following the complementizer; another is preceding the complementizer. These two options both of which are variants of (23) are illustrated in (24):

- (24) a. Su ipa xthes ([tin Eleni]_i) oti [tin Eleni]_i tin idha t_i.
lit. ‘I told you yesterday (Eleni) that Eleni, I saw her’
b. Mislya, ([prestapnika]_i) che [prestapnika]_i sa go hvanali t_i.
lit. ‘I think (the criminal) that the criminal, they got him’

Topic Movement takes place also out of subjunctive complements and indirect questions, as illustrated by the following transformational pairs:

- (25) a. Perimeno [_Tta lefta]_i na ta feri o Jannis t_i → [_TTa lefta]_i perimeno t_i na
ta feri o Jannis t_i.
(I)expect the money to them_{cl.acc} bring Jannis → The money I expect to
bring them_{cl.acc} Jannis
b. Ochakvam [_Tparite]_i da gi donese Ivan t_i. → [_TParite]_i ochakvam t_i da
mi gi donese Ivan t_i.
(I)expect the money to mi them_{cl.acc} bring Ivan → The money I (expect)
to me them_{cl.acc} bring Ivan
- (26) a. Anarotieme [_Tton Janni] jiati i Maria ton paratise.
(I) wonder the Jannis why the Maria him_{cl.acc} left
(Anagnostopoulou 1997: 168, ex. (41b))
b. Chudya se [_TIvan] zashto Mariya go napusna.
(I) wonder Ivan why Maria him_{cl.acc} left

The examples above show that there is position to the left of the subjunctive “complementizers” (modal particles) *da/na*, as well as to the left of the *wh*-word in indirect questions, through which the Topic moves before it continues to the matrix clause.

Hanging Topics are also unboundedly distant from their respective resumptive pronouns. However, differently from CLLD, they cannot appear in any intermediate position (since they are illegitimate in embedded clauses). Consequently, they are not moved from the embedded clause but are directly generated in the matrix clause.

3.1.2. Connectivity effects

In addition to Case connectivity, a second type of connectivity comes from the syntactic behaviour of reflexive pronouns and expressions containing a reflexive pronoun. As is well known, such expressions function as anaphors which need to be bound by their antecedents. In all the Balkan languages under study reflexives are impossible as HTs but are perfectly grammatical as CLLD Topics. Compare the following pairs:

- (27) a. *O eaftos tu_i # o Jannis_i dhen ton frontizi t_i
 himself_{nom} the Jannis not him_{cl.acc} takes-care
 b. Ton eafto tu_i o Jannis dhen ton_i frontizi t_i.
 himself_{acc} the Jannis not him_{cl.acc} takes care
 ‘Jannis doesn’t take care of himself’ (Gr – Anagnostopoulou 1997: 155)
- (28) a. *Cît despre sine_i # Victor_i nu s_i-ar pune in pericol. (Rom – Alboiu 2000: 272)
 ‘As of himself, Victor would not endanger’
 b. Pe sine_i, Victor_i nu s_i-ar pune in pericol t_i.
 ‘Himself Victor would not endanger’
- (29) a. *[Vsichkite si priyateli]_i # gledam da im pomogna t_i (s kakvoto moga).
 (Bul)
 all my(refl) friends (I)try to them_{cl.dat} help with whatever I can
 b. [Na vsichkite si priyateli]_i gledam da im pomogna t_i (s kakvoto moga).
 to all my friends (I)try to them_{cl.dat} help with whatever I can

In the grammatical examples, the anaphor has to reconstruct to its base position (indicated by the trace) in order to be interpreted as bound by its antecedent which shares the same index. The ungrammatical examples, on the other hand, represent a reflexive contained within a HT. Since the anaphor is left unbound, we infer that no reconstruction has taken place. Therefore, such cases constitute evidence that the HT is generated directly in its surface position rather than moved there.

3.2. Island sensitivity

A third piece of evidence which distinguishes between Clitic left dislocation and Hanging Topic left dislocation has to do with so-called ‘islands’. Islands are clauses (or phrases) that do not allow any phrase internal to them to move out. A typical example of islands is an adverbial clause (adjunct clause). The examples below are meant to show that HTs are not sensitive to any islands, because if they were, they would not be able to move out (30a). CLLD, on the other hand, are sensitive to (strong) islands and therefore, movement out of the island is impossible (as indicated in (30b)):

- (30) a. (Kolkoto do) Ivan_i# Mariya se izplashi [_{island} kato go_i vidya kolko zle izglezhda].
 As for Ivan, Mariya simply got-frightened when him_{cl.acc} saw how ill (he)looked
- b. *[Na advokata si]_i Mariya se izplashi, [_{island} sled kato mu_i prizna t_i kakvo e vidyala].
-

*To her lawyer Mariya got frightened after she confessed to him what she had seen

Similar data are reported for Romanian (31) and Greek (Anagnostopoulou 1997: 172):

- (31) a. (Cît despre) Ion_i # am plecat înainte să-l_i-examineze Popescu;
 As for Ion, (I)-have left before to-him_{cl.dat} examine Popescu
 ‘As for Ion, I left before Popescu examined him’
- b. *Pe Ion_i am plecat înainte să-l_i-examineze Popescu
 as for Ion (I)have left before to-him_{cl.acc}-examine Popescu
 (Dobrovie-Sorin 1994: 219)

Table 1. summarizes all the properties of the two types of left dislocation constructions.

Type of construction/ Properties	CLLD	HTLD
1. Case connectivity	YES	NO
2. Tonic pronoun or a clitic pronoun	PRONOMINAL CLTIC	TONIC PRONOUN + PRONOMINAL CLTIC
3. Root or embedded clauses	ROOT AND EMBEDDED CLAUSES	ROOT CLAUSES ONLY
4. Types of phrases	NP, PP, AdvP...	NP ONLY
5. Number of dislocated phrases	MORE THAN ONE	ONE
6. Island sensitivity	SENSITIVE	NOT SENSITIVE

Table 1.

2. Focus constructions

There are several distributional similarities between CLLD Topics and focused phrases in the Balkan languages under study. The syntactic literature on focalization typically distinguishes two types of focus constructions: information focus and left-peripheral focus. The former (IF) is typically realized on the most deeply embedded phrase within the sentence and reports information which the speaker considers discourse-new for his/her interlocutor. In other words, the function of IF is to introduce a new proposition into the Common Ground of the participants in a conversation (Kiss 1998).

The Balkan languages behave standardly when expressing IF as part of the universal inventory of grammatical constraints that languages exploit for discourse purposes: typically, this is the rightmost edge of the sentence with phrasal accent falling on the word or phrase bearing main prominence, as in examples (32)-(30), which employ a standard test for diagnosing the focus bearing constituent (Arnaudova 2010 for Bulgarian; Alboiu 2002; Giurgea 2016 for Romanian; Georgiagentis 2009 for Greek, a.o.).

- (32) A: *Ti efaje o Jannis?* (Gr)
 ‘What did John eat?’
 B: *O Jannis efaje [_{IF}tin turta].*
 ‘John ate Foc[the cake].’ (Georgiagentis 2009, ex.1b)
- (30) A. *Kade hodi tova lyato?* (Bul)
 ‘Where did you go this summer?’
 B. *Hodix do [_{IF}Italiya].*
 ‘I went to Italy’

What is more interesting is that Balkan languages seem to dispose of more structural positions producing different pragmatic interpretations for each focus construction. Consider for example the phenomenon of focus fronting whereby the focused phrase surfaces in the sentence-initial position. Focus fronting is typically employed for expressing emphasis or contrast. Pragmatically, Contrastive Focus (CF) expresses the speaker’s intention to resolve a potential misunderstanding or doubt on the part of his/her interlocutor, or to correct some (part of a previous) statement (this being the reason why some authors working on Romance languages prefer the term ‘corrective’, see Bianchi & Bocci & Cruschina 2015 for Italian).¹⁷ To quote from Zubizarreta (1998): “Contrastive Focus makes a statement about the truth or correctness of (certain aspects of) the presupposition provided by its context statement.” (p. 10). Thus, CF affects the truth conditions of the clause and differently from IF it is often described as ‘narrow’ or ‘unique’.

Unlike Romance, Balkan languages present a far richer inventory of focus constructions and moreover, interpretationally fronted (preverbal) focus need not

be of the contrastive or corrective type (see Giurgea 2016 for Romanian; Krapova 2021 for Bulgarian) even though this type of focus is also common and can be accommodated to English cleft constructions, which are missing in all Balkan languages:

- (31) A: I-a spus lui Ion
 ‘He told Ion’
 B: Nu, **lui George** i-a spus
 no dat George him_{cl.dat}-has told
 ‘No, he told **George**/ It was **George** he told’ (Rom – Giurgea 2016, ex. 44))

Sentence-initial positions are exploited for another type of focus which too bears some resemblance to English clefts but is more widespread than clefting since the fronted phrase need not be accompanied by a focusing particle (like ‘only’ or additives like ‘and’ and ‘also’). Both definite and ‘bare’ phrases, i.e., unaccompanied by any definite or indefinite determiner, can appear in the sentence-initial position which counts for the purposes for focalization (see also Nicolova 2000, Ницолова / Nitsolova 2008: 150):

- (32) a. **Samo cvetya/ cvetyata** shte ì podarya/ **I edno cvete** dazhe ne mi donese.
 (Bul)
 only flowers/the flowers will I give her/ not even a flower did he bring me
 b. **Mono ta luludhia** dialekxa moni mou; **Ke luludhia** aghorasa. (Gr)
 only the flowers (I)chose by myself; and flowers (I)bought
 c. **Și pe Maria** am anunțat-o (Rom)
 also pe Maria (I)have informed-her_{cl.acc}
 ‘I also informed Maria’

A third type of focus realized in all Balkan languages is what some authors label ‘mirative’ (Bianchi & Bocci & Cruschina 2016). Mirative focus carries an evaluative import since it typically expresses unexpected or surprising information that is not yet part of the system of beliefs of the interlocutors (Cruschina 2012). This type of focus does not rely on a salient proposition in the discourse and does not necessarily introduce a set of potential alternatives to the focused phrase, unlike the mirative focus of Romance languages:

- (33) a. **O mie de lei** am dat pe aparatul ăsta!
 a thousand of lei (I)have given on device-the this
 ‘I paid a thousand lei for this device!’ (Giurgea 2016, ex.2)
 b. Predstavyash li si? **V Alyaska** xodil!
 ‘Can you imagine? To Alaska he went!’

c. **Golyam glupak** si bil!

You are such a fool!'

In spontaneous (colloquial) speech, focus is also used in answers to *wh*-questions (which introduce an open set of alternatives) as well as in answers to polar questions (which introduce a limited set of alternatives out of which only one is selected as the actual alternative for which the predicate holds), cf. (34):

(34) a. Koya godina stana tova?

Mayche **2004** beshe.

b. Za London li zamina Ivan?

Ne, **za Berlin** zamina.

c. Ion s-a dus la Londra?

(Rom)

Ion REFL-has gone to London

Nu, **la Berlin** s-a dus (nu la Londra).

No, to Berlin REFL-has gone not to London

Given that *wh*-questions introduce an open set of alternatives, it can be summarized that in order for a focused constituent to be fronted in the Balkan languages, it can either be discourse-new (34a,b), emphatic (32) or mirative (33), with a host of specific interpretations whose intonational contour deserves a more detailed study. Giurgea (2016) argues for Romanian that in cases comparable to (34c) the fronted focus position is marked by the feature [exclusive] (i.e., it is neither mirative nor exclamative, nor introduced by a focal particle) and introduces (i) the presupposition that at least one alternative is true and (ii) an exhaustivity inference, which can be treated as a presupposition: if the proposition at hand *p* is true, any other focal alternative (not entailed by *p*) is false.

Another common property of the Balkan focus constructions is the multiplicity of positions where a focus-marked element can appear: a) displaced in a matrix clause even though it belongs to an embedded clause; b) in front of a declarative or an interrogative complementizer; c) in front of a *wh*-word/phrase in a *wh*-question.¹⁸ All of these properties are attested in Greek (35) and Bulgarian (36):

(35) a. [_F **Ti Maria**] lene oti pandreftike o Jannis (Gr – Tsimpli 1995: 193).
the_{acc} Maria (they)say that married the_{nom} Jannisb. Lene [_F **ti Maria**] oti pandreftike o Jannis
(they)say the Maria_{acc} that he married the_{nom} Jannis
(Joseph & Philippaki-Warburton 1987: 104)c. Me rotise [_F **ta vivlia**] pjos aghorasa (Alexiadou 1997: 73).
me_{acc} (he/she)asked the books who bought

- (36) a. [_F **Mariya**] mislya, che shte izberat za predsedatel. (Bul)
 Maria (I)think that will (they) elect for chairman’
- b. Ivan znaex, che shte xodi, no [_F **ti**] che shte xodish, ne znaex
 Ivan (I) knew that will (he)go but you that will (you)go, not (I)knew
 ‘I knew that Ivan would go but I didn’t know that you would go too’
 (Rudin 1991)
- c. Chudim se [_F **na svekara**] kakvo da podarim.
 (we)wonder to father-in-law what to give-as-present
 ‘We are wondering what present to give to my father-in-law’

Given that these properties instantiate a movement derivation involving the left periphery of the clause, we can conclude that the dislocated position of Focus is also derived by movement: (35a)/(36a) show instances of unbounded (long-distance) Focus movement; (35b)/(36b) show instances of short Focus movement (to the left periphery of an embedded declarative complement; (35c)/(36c) show instances of short Focus movement in embedded *wh*-questions.¹⁹

3. Clitic doubling of fronted constituents

Given the data discussed so far, we can generalize that in Greek and in Bulgarian, constituents that can be topicalized are also eligible for focalization. In other words, as predicted by the abstract structures in (4) above, the two constructions should be syntactically differentiated through the presence vs. absence of a resumptive clitic (in the case of object noun phrases). This, however, does not seem to be the case in Romanian. As reported by Dobrovie-Sorin (1990), Cornilescu (2000), and Alboiu (2000), in this language not just Topics but also focused phrases can be clitic resumed:

- (37) a. **Pe** [_F **Petru**] Maria nu l-ar ajuta, pe Gheorghe, da
 OBJ Petre Maria not him helps, OBJ George yes
 ‘It’s Peter that Maria doesn’t help, she helps George’
 (Rom- Dobrovie-Sorin 1990: 220)
- b. Eu [_F **pe Popescu**] l-am vazut (nu pe Ionescu)
 I pe Popescu him_{cl,acc}-have read (not OBJ Ionescu)
 ‘I read Popescu (not Ionescu)’

While all authors acknowledge that Romanian observes the pan-Balkan ban on doubling of ‘bare’ nouns (i.e. nouns without any determiner), Dobrovie-Sorin (1994) and Cornilescu (2000) nevertheless give examples of focused definite phrases

where clitic resumption is not just possible but obligatory even in the presence of focus particles (like *numai* ‘only’, *chiar* ‘even’, *macar* ‘at least’):²⁰

- (38) a. **Numai pe Ion il** iubește Maria. (Rom)
 only OBJ Ion him_{cl.acc} loves Maria
 b. **Macar cartea asta** au citit-o elevii.
 at least book this have read-it_{cl.acc} students

The Bulgarian and Greek equivalents of (38) are ungrammatical, as (39) shows:

- (39) a. ***Samo Ivan go** obicha Mariya (Bul)
 only Ivan him loves Mariya
 b. ***Mono ton Jani ton** agapai i Maria. (Gr)
 only Jannis him loves Maria

As mentioned above, we suggest more generally, that wherever there are differences between the three languages, these seem to be determined by independent language-internal properties.

One could think that the contrast between (33) and (34) is a primitive, i.e. non-derived, difference between Romanian and Bulgarian/Greek. But this may well turn out to be related to an independent difference between these languages, namely to the fact that Romanian is not as restricted as Bulgarian and Greek in its use of true clitic doubling (*anticipatio*) where the NP double remains in situ rather than moving to the left periphery. See the contrast in (40):

- (40) a. L-am văzut numai pe Ion. (Rom)
 him_{cl.acc}-have seen only OBJ Ion
 ‘I saw only Ion’
 b. *Az go vidyax samo Ivan. (Bul)
 I him_{cl.acc} saw only Ivan

Whatever the explanation for the distribution of clitic resumed phrases in Romanian focus construction, it is tempting to say that the contrast in (40) is at the basis of that between (38)-(39), if we presume that the clitic doubled noun phrase originates in a postverbal position and then moves to a preverbal position without further changes in the structure. This is rendered plausible by the following two facts:

1) Whenever clitic doubling is impossible in Romanian (as with indefinite quantifiers like *pe altcineva* ‘someone else’ illustrated in (41))²¹, resumption of the same phrase in preverbal focus is also impossible (which makes one think that the

object preposition *pe* is a necessary but not a sufficient condition for clitic doubling (Dobrovie-Sorin 1994: 220ff; Hill & Tasmowski 2008):

- (41) a. *Ion l-așteaptă pe altcineva vs. Ion așteaptă pe altcineva.
 Ion him_{cl.acc}-waits(for) OBJ someone else'
 b. *Ion **pe altcineva** l-așteaptă, nu pe Maria vs. Ion **pe altcineva**
 așteaptă, nu pe Maria
 'Ion OBJ someone else waits(for) not OBJ Maria'

2) Whenever clitic doubling qua *anticipatio* is obligatory in Bulgarian or Greek (as happens with psychological predicates), the fronted focus phrase must also be clitic resumed (Krapova & Cinque 2008):

- (42) a. Boli go glavata Ivan (Cf. *Boli glavata Ivan)
 hurts him_{cl.acc} head-the Ivan hurts head-the Ivan
 b. (Samo) **Ivan** go boli glavata. (Cf. *(Samo) **Ivan** boli glavata)
 only Ivan him_{cl.acc} hurts head-the only Ivan hurts head-the

4. Linear orders of Topic and Focus in the Balkan languages

Finally, another property shared by all of the Balkan languages under study is the relative order of Topics and Focus in the left periphery. In a single clause, there can be multiple CLLD Topics but there is always a single Focus per clause (also known as 'Focus uniqueness requirement'). Moreover, in conformity to the universal organization of the Left Periphery (Rizzi 1997), Topics must precede all phrases that can be argued to possess a focus feature (Horvath 1986): contrastively focused phrases, bare quantifiers, as well as *wh*-phrases. There is also a tendency for these latter constituents to appear adjacent to the verbal predicate. Examples (from Krapova 2004) are provided below from the three languages under study:

- (43) a. [_TMariei] [_Tflorile acestea] tu nu i le poți cumpăra.
 For Maria these flowers you cannot buy her them
 (Rom – Cornilescu 2001)
 b. [_TMariei] [_F**flori**] este potrivit să-i oferi.
 To Maria flowers is appropriate to give
 c. [_TPe Victor] [_F**cine**]-l așteaptă la aeroport. (Alboiu 2000: 260, ex. (28a))
 OBJ Victor who-him_{cl.acc} is going to wait at the airport
 'Who is going to wait for Victor at the airport?'

- (44) a. [_TNa Mariya] [_T tezi cvetya] sam ì gi podaril az. (Bul)
 To Mariya these flowers have her._{cl.dat} them._{cl.acc} given I
 b. [_T Mariya] [_F **mazhat i**] ya izvika i tya se pribra.
 Mariya, husband-the her her._{cl.:acc} called and she left home
 c. [_T I nego] [_F **koy**] go pita, ama na – kato e za razvala, i toy e tam.
 And him who him._{cl.acc} asks, but when it comes to problems, he is there too
- (45) a. [_TTa vivlia] [_F **sti Maria**] ta edhosa. (Gr – Alexiadou 1997: 74)
 The books to Maria them._{cl.acc} (I)gave
 b. Me rotise [_T sti Maria] [_F **pjos**] tis edhose afta ta vivlia.
 me._{cl.acc} asked to Maria who her._{cl.dat} gave these the books

Based on all the comparative data discussed above, regarding both topicalization and focalization structures, we can summarize the overall order of dislocated phrases in the left periphery of the three languages under study (Greek, Bulgarian and Romanian) with the following structural hierarchy:

(46) *as for* phrases > hanging topics > clitic left dislocated topics (Aboutness Topic) > clitic left dislocated topics (Given Topic(s)) > focalized phrases

5. Conclusion

The organization of the Left Periphery in the Balkan languages, including the relative order of Topic and Focus, reflects a stable typological tendency rather than a pure *Sprachbund* effect. Nevertheless, the development of the common discourse patterns can be seen as a follow-up process on some of the convergence phenomena (object reduplication and the morpho-syntactic expression of definiteness), which, among other phenomena lead to the establishment of the Balkan Language Union (Асенова / Asenova 2002). According to Минчева / Mincheva (1969), Topic structures illustrate some of the most specific properties of the syntax of colloquial speech: shaping of intonational-syntactic groups, the possibility for segmentation of the utterance which “deviates” from the norms of the standard language, ellipsis, pleonasm, etc. These principles have manifested themselves at quite an early stage in the Balkan context. The same could be hypothesized for Focus structures which not only allowed for the independent syntactic expression of (different kinds of) non-presupposed information, but also create additional stylistic effects. Given the colloquial nature of the bi- and multi-linguistic contacts at the time when the main Balkanisms were integrated into the structure of each language, the universal principles of (colloquial) syntax must have fed the general Balkan tendency towards a greater word order freedom.

While the existence of left peripheral structures is by no means an original Balkan phenomenon,²² the Balkan left periphery is worth studying at least for two reasons. On the one hand, such a study can offer support for the organization of that

particular area in the clause specifically dedicated to the expression of discourse relevant concepts, especially in the languages of the Balkan *Sprachbund*, which are characterized by high ‘discourse prominence’. On the other hand, given that even in the modern languages Topic and Focus structures are most typical for colloquial speech, one could hypothesize that the quasi-identical ordering of phrases in sentence initial positions can find its explanation in the mechanisms underlying the need for efficient communication: structures with overtly marked grammatical or discourse functions are often favoured by speakers in bi- or multi-lingual contact situations (Lindstedt 2000). Thus, the emergence of a well-defined left periphery in the Balkan clause could have been facilitated by the independent emergence, in different stages of the different languages, of discourse strategies related to saliency, accessibility, prominence or familiarity which in turn gave rise to explicit marking of topic and focus structures through shared syntax for the purpose of achieving greater communicative efficiency. While historical considerations have not played a role in the present contribution, this is nevertheless a direction worth being explored in the future.

NOTES

¹ We follow Лопашов / Lopashov (1978, 14) in differentiating two types of structures: those in which the object is preposed with respect to the verb (*reprisa*), and those in which the object is postposed (*anticipatio*). Although he considers the difference in quantitative terms, there are other, deeper, differences between these two structures. There are also historical considerations for such a distinction, at least in Bulgarian. As reported by Минчева / Mincheva (1969), Topicalization qua preposing of the object is a much older phenomenon, as it can be found in several contexts in Old Bulgarian. Typically, the anaphoric pronoun or a demonstrative pronoun used to double the preposed (heavy and intonationally independent) object. *Anticipatio*, on the other hand, is a later phenomenon – the earliest documents in which it is attested date from the 12th – 13th c. According to Минчева / Mincheva, the later expansion of *anticipatio*, while still attributable to the syntactic principles of colloquial speech, involves additional factors such as the position of the enclitic, the syntactic independence of the verbal group, etc.

² In all the examples to follow, focused phrases will be given in boldface.

³ In Romanian, clitic doubled direct objects are usually introduced by the marker *pe* if referring to a human definite constituent or a proper name. While *pe* is standardly analyzed as a preposition (Dobrovie-Sorin 1994), some authors view it as a pragmatic marker (Hill & Tasmowski 2008) whose function is to increase the prominence of the direct object’s referent by “highlighting the direct object against a background set and locating it in the speaker’s perspective” (p.145).

⁴ Historically, the primary function of the *reprisa* has been related to the grammaticalization of the SVO word order in the Balkan languages, following the loss of Case distinctions, whose most visible effects are observed in Bulgarian. Apart from ensuring a greater word order freedom and achieving discourse prominence, the topicalization of the object in a sentence initial position serves other syntactic purposes, such as the disambiguation of (potentially ambiguous) subject – object structures (cf. Лопашов / Lopashov 1978: 83, 99, 101-105; Асенова / Asenova 2002: 108f), e.g. *Dimov go ubi Meri Lamour* lit. ‘Dimov him killed M.L. (ex. from Попов / Popov 1962).

⁵ This type of Focus should be strictly differentiated from New Information Focus, which, as mentioned above, corresponds best to the traditional notion of *rheme* and appears in a sentence final position, since it can be used as an answer to a request for new information, e.g. *Kakvo donese Ivan?* ‘What did Ivan bring?’- *Ivan donese* [_{IF} knigite] ‘Ivan brought the books’.

⁶ The two structures differ in terms of the operator status of the moved element which can bind a variable in (4b) as focalized constituents count as operators, while topicalized ones do not, whence the presence of a clitic which binds the trace more locally (see for discussion Cinque 1977, 1990).

⁷ In Greek, dislocation for emphasis can be accompanied by an emphatic nonclitic proform, cf. *Stin Elada, eki na pame jia djakopes* lit. ‘In Greece, there let’s go on vacation’; *Tin kiriaki, tote na pame* lit. ‘Sunday, then let’s go’ (Joseph & Philippaki-Warburton 1987: 100)

⁸ Especially if it performs the grammatical function of a direct object (for a discussion see Hill & Tasmowski 2008; Krapova & Cinque 2008). This condition is relaxed for indirect objects.

⁹ In discussing cases like (6) in Bulgarian, Guentchéva (1994) argues that the extraposed term (the HT in our terminology) is co-referential only with (and reduplicated only by) the tonic pronoun. She gives as an example the sentence *A sarceto bjas go kasa nego kleto* lit. ‘And the poor heart, anger tears it apart’ (p. 157) noting that if the initial topic is set off by a pause, the presence of a tonic pronoun crucially distinguishes the case of extraposition from the regular case of clause-internal clitic doubling. Since this is true for Romanian and Greek, we can conclude that no Balkan language admits a sequence of a case-marked lexical NP, a clitic and an emphatic pronoun within the same clause, i.e., “triple reduplication” is inexistent.

¹⁰ Some authors (Dobrovie-Sorin 1994; Anagnostopoulou 1997) working on the Balkan languages use the term **Left Dislocation (LD)** to refer to what we label here HTLD. It should be noted however that LD is different from so-called **English Left Dislocation**, which at least in some of its instantiations may introduce discourse-new material and does not always require an anaphoric element inside the clause (for a discussion see Prince 1988).

¹¹ Following the standard practice, in the examples below the (heavy) intonational pause after the HT will be indicated with the symbol #

(i) We went to Florida last summer, and we went to Disney World. The best ride the whole time was Jurassic Park. It was so scary. *My sister Chrissie, her eyes* were poppin’ out. (CL 10/14/00)

(ii) (As for) *flowers*, I like *tulips*.

¹² The HTLD construction can feature an epithet like *gliko mu* ‘my sweetie’ in (9). This is true also for the other Balkan languages.

¹³ According to Bertolussi (2017), in Latin *Nominativus pendens* is the most popular type of dislocation and constitutes the typical form of Hanging Topic.

¹⁴ We do not mean that all criteria need to be met in order for a certain construction to qualify as HTLD. For example, if a tonic pronoun is not realized in a certain topic structure, then Case connectivity becomes the distinguishing factor with respect to CLLD. Compare the two corpus examples taken from Джонова / Dzhonova (2004): *Na Ivan otdavna ne sa mu plashtali* lit. ‘To John, they haven’t paid him for a long time’ vs. *Ivan otdavna ne sa mu plashtali* lit. ‘John, they haven’t paid him for a long time’.

¹⁵ The tonic pronoun can also occur at the absolute end of the sentence. In this case, we are dealing with the mirror image of the CLLD – Clitic Right Dislocation (CLRD). The latter is also typical for marking a Topical object (or as a kind of an afterthought) in Bulgarian, as well as in the other Balkan languages (Tisheva 2007).

¹⁶ An anonymous reviewer points out that the same prosodic effect is achieved with the interrogative particle *li*, especially with first person nominative topics, e.g., *Az li, na men tova nikoga ne mi se e sluchvalo*. 'Talking about me/as for me, this never happened to me'. Here, *li* is a not a yes-no question marker but performs a pragmatic function highlighting the topic.

¹⁷ Syntactically, this strategy makes use of the *subjective word order* ('obraten slovored', cf. Ivančev 1978): the information is presented as the most relevant part of the utterance and is typically pronounced with (strong) emphasis, i.e., it carries emphatic stress ('logichesko udarenie', cf. Попов / Попов 1961; Цыхун / Сухун 1962: 287). This type of Focus fronting operation conveys new information only indirectly: by emphasizing the information the speaker typically brings forward a (potentially) novel quality or property of what is being talked about, i.e., about the discourse theme.

¹⁸ This co-occurrence is not possible in matrix clauses, probably for independent reasons.

¹⁹ This last possibility is also attested in Romanian, according to Cornilescu (2000) who cites cases like (i):

- (i) Nu știu **alții** cum sunt, dar **eu** îmi aduc aminte de asta cu plăcere.
'I don't know how **others** are but **I** remember this with pleasure'

²⁰ According to Cornilescu (2000), in these examples doubling is obligatory because of the inherent semantics of the proper names or of the definite descriptions, which are "good" topics.

²¹ According to Alboiu (2000), certain quantifiers (both universal and distributive) like *oricine* 'anyone', *fiecare* 'each' can be clitic resumed, as opposed to 'bare' quantifiers like *fiece* 'every', and *cineva* 'someone'. The author argues that, depending on their inherent semantics, quantifiers behave as CLLD Topics or as Focus. Hence their split behaviour.

²² A fine-grained left periphery has been postulated for a wide range of language groups in numerous cross-linguistic studies (for references to concrete works, see Rizzi & Bocci 2017).

REFERENCES

- Alboiu 2000: Alboiu, G. *The features of movement in Romanian*. Doctoral dissertation. University of Manitoba.
- Alexidaou 1997: Alexiadou, A. *Adverb Placement. A case study in antisymmetric syntax*. Amsterdam: John Benjamins.
- Anagnostopoulou 1994: Anagnostopoulou, E. *Clitic dependencies in Modern Greek*. Doctoral dissertation. University of Salzburg.
- Anagnostopoulou 1997: Anagnostopoulou, E. Conditions on clitic doubling in Greek. In: Van Riemsdijk, H. (Ed.): *Clitics in the languages of Europe*. Berlin: Mouton De Gruyter, 761-798.
- Arnaudova 2001: Arnaudova, O. Prosodic movement and information focus in Bulgarian. In: Franks, S., King, T. H., Yadroff, M. (Eds.): *Annual Workshop on Formal*

Approaches to Slavic Linguistics. The Bloomington Meeting 2000. Ann Arbor: Michigan Slavic Publications, 19-36.

Arnaudova 2010: Arnaudova, O. *Focus and Bulgarian Clause Structure: Word Order Variation and Prosody.* Riga: VDM Verlag Dr. Müller.

Bally 1932/1965: Bally, Ch. *Linguistique générale et linguistique française.* Paris: Ernest Leroux, Vendôme.

Bianchi & Cruschina 2016: Bianchi, V. & Cruschina, S. The derivation and interpretation of polar questions with a fronted focus. – *Lingua* 170 (2016), 1, 47-68.

Bianchi & Frascarelli 2010: Bianchi, V. & Frascarelli, M. Is topic a root phenomenon? – *Iberia* 2 (2010), 43-88.

Bianchi & Bocci & Cruschina: 2015 Bianchi, V., Bocci, G., Cruschina, S. Focus fronting and its implicatures. In: Aboh, E., Schaeffer, J., Sleeman, P. (Eds.): *Romance Languages and Linguistic Theory* 2013. Amsterdam: John Benjamins, 1-20.

Bianchi & Bocci & Crischina 2016: Bianchi, V., Bocci, G., Cruschina, S. Focus fronting, unexpectedness, and the evaluative dimension – *Semantics and Pragmatics*, 9, 1-54.

Bertolussi: 2017: Bertolussi, B. Topicalizations, left dislocations and the left periphery – *Catalan Journal of Linguistics* 16 (2017), 101-123.

Chomsky 1973: Chomsky, N. Conditions on transformations. In: Anderson, S., Kiparsky, P. (Eds.): *A Festschrift for Morris Halle*, New York: Holt, Rinehart & Winston, 232-286.

Cinque 1977: Cinque, G. The movement nature of left dislocation – *Linguistic Inquiry* 8 (1977), 397-412.

Cinque 1990: Cinque, G. *Types Of A' Dependencies.* Cambridge, Mass.: MIT Press.

Cornilescu 2000: Cornilescu, A. Rhematic focus at the left periphery: the case of Romanian. Paper presented at *Going Romance*, Utrecht.

Cornilescu 2001: Cornilescu, A. *On focusing and wh-movement in Romanian.* Ms., University of Bucharest.

Cruschina 2012: Cruschina, S. *Discourse-Related Features and Functional Projections.* Oxford & New York: Oxford University Press.

Dobrovie-Sorin 1990: Dobrovie-Sorin, C. Clitic doubling, wh-movement, and quantification in Romanian – *Linguistic Inquiry*, 21 (1990), 3, 351-398.

Dobrovie-Sorin 1994: Dobrovie-Sorin, C. *The Syntax of Romanian.* Berlin: Mouton De Gruyter.

Georgiafentis 2009: Georgiafentis, M. Focus marking in Greek: syntax or phonology? In: Jiannakis, G., Baltadzani, M., Ksidopoulos, G., Tsangalidis, A. (Eds.): *Proceedings of the 8th International Conference on Greek Linguistics.* Thessaloniki: Aristotle University of Thessaloniki, 2009, 621-636.

Giurgea 2016: Giurgea, I. On the interpretation of focus fronting in Romanian – *Bucharest Working Papers in Linguistics*, 18 (2), 37-61.

Guentchéva 1994: Guentchéva Z. *Thématisation de l'objet en bulgare.* Bern: Peter Lang S.A.

Horvath 1986: Horvath, J. *Focus in the Theory of Grammar and the Syntax of Hungarian.* Dordrecht: Foris.

Joseph & Philippaki-Warbuton 1987: Joseph, B., Philippaki-Warbuton, I. *Modern Greek*. London: Croom Helm.

Kiss 1998 : Kiss, K. É. Identificational focus versus information focus – *Language* 74 (1998), 245-273.

Krapova 2002: Krapova, I. On the left periphery of the Bulgarian sentence. – *University Of Venice Working Papers in Linguistics*, vol. 12, 2002.

Krapova 2004: Krapova, I. Word order in topic-focus structures in the Balkan languages. In: Basciani, A., Tarantino, A. (Eds.): *L'Europa d'oltremare. Contributi italiani al IX congresso Internazionale dell' Association Internationale d'Etudes du Sud-Est Europeen, Tirana, 30 agosto – 3 settembre 2004. Romania Orientale* 7, XVII. Roma: Bagatto Libri. 139-16.

Krapova 2021: Krapova, I. Complementizers and particles inside and outside of the Bulgarian left periphery. In: Wiemer, B, Sonnehauser, B. (Eds.): *Clausal Complementation in South Slavic*. Berlin: Mouton, 211-269. (Trends in Linguistics. Studies and Monographs. Vol. 361.)

Krapova & Cinque 2005: Krapova, I., Cinque, G. Two asymmetries between clitic left and clitic right dislocation in Bulgarian. In: Broekhuis, H., Corver, N., Huybregts, R. Kleinhenz, U., Koster, J. (Eds.): *Organizing Grammar. Linguistic Studies in Honor of Henk van Riemsdijk*. Berlin: Mouton de Gruyter, 359-364.

Krapova, & Cinque 2008: Krapova, I., Cinque, G. Clitic Reduplication constructions in Bulgarian. In: Kallulli, D., Tasmowski, L (Eds.): *Clitic Doubling in the Balkan Languages*. Amsterdam/Philadelphia: John Benjamins, 257-288.

Krifka 2001: Krifka, M. For a structured account of questions and answers. In: Féry, C., Sternefeld, W. (Eds.): *Audiatur vox sapientiae. A Festschrift for Achim von Stechow*. Berlin: Akademie-Verlag, 287-319.

Lindstedt 2000: Lindstedt, J. Linguistic balkanization: contact-induced change by mutual reinforcement. In: Gilbers, D., Nerbonne, J., Schaeken, J. (Eds.): *Languages in Contact*. Amsterdam & Atlanta, Ga: Rodopi, 231-246. (Studies in Slavic and General Linguistics, 28.)

Nicolova 2000: Nicolova, R. Kontrastfokussierung mit Partikeln im Bulgarischen. – *Linguistische Arbeitsberichte* 75, 103-115.

Ott & de Vries 2014: Ott, D., de Vries, M. A biclausal analysis of right dslocation. In: Huang, H.-L, Poole, E., Rysling, A. (Eds.): *Proceedings of NELS* 43, Vol. 2, Amherst: GLSA, 41-54.

Philippaki-Warbuton et al.; 2004: Philippaki-Warbuton, I., Varlakosta, S, Georgiafentis, M., Kotzoglou, G. Movingf from theta-positions: pronominal clitic doubling in Greek – *Lingua* 114 (2004), 963-989.

Prince 1998: Prince, E. F. On the limits of syntax, with reference to left-dislocation and topicalization. In: Cullicover, P., McNally, L. (Eds.): *The Limits of Syntax. Syntax and Semantics* 29. New York: Academic Press, 261-302.

Rizzi 1997: Rizzi, L. The fine structure of the left periphery. In: Haegeman, L. (Ed.): *Elements of Grammar*. Dordrecht: Kluwer, 281-337.

Rizzi & Bocci 2017: Rizzi, L. Bocci, G. Left periphery of the clause: primarily illustrated for Italian. In: Everaert, M., van Riemsdijk, H. (Eds): *The Wiley Blackwell Companion to*

Syntax (2nd edition). Hoboken, NJ: John Wiley and Sons, 2171-2200. <<https://onlinelibrary.wiley.com/doi/epdf/10.1002/9781118358733.wbsyncom104>>

Rudin 1986: Rudin, C. *Aspects of Bulgarian Syntax: Complementizers and Wh-Constructions*. Columbus, Ohio: Slavica Publishers.

Rudin 1991: Rudin, C. Topic and focus in Bulgarian – *Acta Linguistica Academiae Scientiarum Hungaricae*, vol. 40 (1991), 3-4, 429-447.

Samek-Lodovici 2015: Samek-Lodovici, V. *The Interaction of Focus and Givenness*. Oxford: Oxford University Press.

Schwarzschild 1999: Schwarzschild, R. Givenness, Avoid F and other constraints on the placement of accent – *Natural Language Semantics* 7, 141-177.

Tisheva 2007: Tisheva, Y. On right dislocation and marginalization in Bulgarian. In: Kosta, P, Schürcks, L. (Eds.): *Linguistics Investigations into Formal Description of Slavic Languages: Contributions of the Sixth European Conference held at Potsdam University November 30 – December 2, 2005*. Frankfurt am Main: P. Lang, 389-397.

Tsimpli 1995: Tsimpli, J.-M. Focusing in Modern Greek. In: K. E. Kiss (Ed.): *Discourse-Configurational Languages*. Oxford: Oxford University Press, 176-206.

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

Vallduvi 1992: Vallduvi, E. *The informational Component*. New York: Garland.

Zubizarreta 1998: Zubizarreta, M.-L. *Prosody, Focus and Word Order*. Cambridge, Mass.: MIT Press

АГ 1994: Академична граматика на съвременния български книжовен език. Том 3. Синтаксис. София: БАН. [AG 1994: *Akademichna gramatika na savremenniya balgarski knizhoven ezik*. Tom 3. Sintaksis. Sofia: BAN.].

Асенова 2002: Асенова, П. *Балканско езикознание. Основни проблеми на Балканския езиков съюз*. Велико Търново: Фабер. [Asenova 2002: Asenova, P. *Balkansko ezikoznanie. Osnovni problemi na Balkanskiya ezikov sayuz*. Veliko Tarnovo: Faber.].

Джонова 2004: Джонова, М. Конструкции от типа *аз ми се спи* в българската разговорна реч. В: *ged. Проблеми на българската разговорна реч*. Кн. 6. Велико Търново: Университетско издателство „Св. св. Кирил и Методий“, 107-116. [Dzhonova 2004: Dzhonova, M. Konstruktsii ot tipa *az mi se spi* v balgarskata razgovorna rech. V: *Problemi na balgarskata razgovorna rech*. Vol. 6. Veliko Tarnovo: Universitetsko izdatelstvo “Sv.sv. Kiril i Metodiy”, 107-116.].

Иванчев 1978: Иванчев, С. *Приноси в българското и славянското езикознание*. София: Наука и изкуство. [Ivanchev 1978: Ivanchev, S. *Prinosi v balgarskoto i slavyanskoto ezikoznanie*. Sofia: Nauka i izkustvo.].

Лопашов 1978: Лопашов, Ю. А. *Местоименные повторы дополнения в балканских языках*. Ленинград. [Lopashov 1978: Lopashov, Ju. A. *Mestoimennyye povtory dopoleniya v balkanskix jazykax*. Leningrad.].

Минчева 1969: Минчева, А. Опыт за интерпретация на модела на удвоените допълнения в българския език. – *Известия на Института за български език XVII* (1969), 3-50. [Mincheva 1969: Mincheva, A. Opit za interpretatsiya na modela na udvoenite dopalneniya v balgarskiya ezik – *Izvestiya na Instituta za balgarski ezik XVII* (1969), 3-50.].

Ницолова 2008: Ницолова, Р. *Българска граматика. Морфология*. София: Университетско издателство „Св. Климент Охридски“. [Nitsolova 2008: Nitsolova, R. *Balgarska gramatika. Morfologiya*. Sofia: Universitetsko izdatelstvo “Sv. Kliment Ohridsky”].

Попов 1961: Попов, К. О логическом ударении. – *Вопросы языкознания* 1961, 3, pp. [Popov, K. O logicheskom udarenii – *Voprosy jazykoznanija* 1961, 3, pp.].

Попов 1962: Стилно-граматична употреба на удвоеното допълнение в българския книжовен език – *Известия на Института за български език XVIII* (1962), 459-470. [Popov 1962: Popov, K. Stilno-gramatichna upotreba na udvoenoto dopalnenie v balgarskiya knizhoven ezik – *Izvestiya na Instituta za balgarski ezik XVIII* (1962), 459-470.].

Цыхун 1962: Цыхун, Г. Местоименната клитика и словоредът в българското изречение – *Български език XII* (1962), 4, 283-291. [Сухун 1962: Сухун, Г. Mestoimennata klitika i slovoredat v balgarskoto izrechenie – *Balgarski ezik XII* (1962), 4, 283-291.].

Цыхун 1981: Цыхун, Г. *Типологические проблемы балканославянского языкового ареала*. Москва: Наука и техника. [Сухун 1981: Сухун, Г. *Tipologicheskie problemy balkanoslavjanskogo jazykovogo areala*. Moskva: Nauka i tehnika.].

Assoc. Prof. Piyana Krapova

Ca' Foscari University of Venice

Venice, Italy

e-mail: krapova@unive.it