John Benjamins Publishing Company

The relative order of foci and polarity complementizers

A Slavic perspective

Elena Callegari
University of Iceland

According to Rizzi & Bocci's (2017) suggested hierarchy of the left periphery, fronted foci (FOC) can never precede polarity complementizers (PolC); yet languages like Bulgarian and Macedonian appear to display precisely such an ordering configuration. On the basis of a cross-linguistic comparison of ten Slavic languages, I argue that in the Slavic subgroup the possibility of having a focus precede PolC is dependent on the morphological properties of the complementizer itself: in languages where the order FOC < PolC is acceptable, PolC is a complex morpheme derived through the incorporation of a lower functional head with a higher one. The order FOC < PolC is then derived by giving overt spell-out to the intermediate copy of PolC rather than to the topmost one. In turn, this option is linked to the possibility, recorded in all languages which allow for FOC < PolC, to also realize the morpheme expressing interrogative polarity as an enclitic particle attaching to fronted foci.

Keywords: left periphery, fronted foci, polarity questions, complementizers, slavic, word order

1. Introduction

In this article, I will be concerned with accounting for cross-linguistic variation in the relative distribution of two types of left-peripheral elements. The first are polarity complementizers (PolC), namely complementizers whose function is that of introducing embedded polarity questions; the second are fronted types of constituents in narrow focus. I provide an example of a configuration containing both elements in (1), from Italian; in (1), the polarity complementizer *se* (= 'if') is marked in bold, whereas the fronted focus -in this case, a PP- is in capitals.

(1) *Mi   domando **se** A   TROMSØ Espen sia   nato*   (ITA)
     Refl. I-wonder **if** IN TROMSØ Espen is(subv.) born
     'I wonder if Espen was born IN TROMSØ'

Concerning the relative order of these two elements, a constituent in focus will either follow the polarity complementizer, or it will precede it; I refer to the former configuration as *alpha* (2a), and to the latter as *beta* (2b). Note that I use the symbol '<' to indicate linear precedence:

(2) a.   Order α: if < FOC (**se** A TROMSØ)
     b.   Order β: FOC < if (A TROMSØ **se**)

According to Rizzi (2001) (see also Rizzi 2004, and Rizzi & Bocci 2017), both foci and polarity complementizers are associated with dedicated left-peripheral projections. In particular, the complementizer 'if' is argued to be externally merged in the head of Int(errogative)P, a projection sandwiched in between two other projections equally specialized for hosting complementizers, *Force* and *Fin*. Rizzi's suggested derivation for fronted foci is slightly more complex, since these are not taken to be base-generated directly in the left periphery. Rather, they are first externally merged in their argumental position, where they can receive case and theta role. Only at a second stage are they moved to the left periphery, where they surface in the specifier of a dedicated *Focus* phrase. This dedicated FocP is non-recursive and is lower than IntP, as illustrated in (3). (3) represents the most updated[1] version of the cartographic hierarchy of the left periphery:

(3) Force < TopP < IntP < TopP < FocP < TopP < ModP< TopP < QEmb < Fin < IP

<div align="right">(Rizzi & Bocci 2017, Example (29))</div>

Evidence in favor of the idea that FocP must be lower than IntP comes from grammaticality patterns such as the ones illustrated in (4). The pair in (4) shows that, in languages like Italian – on which (3) is primarily based–, a constituent in narrow focus may only follow the polarity complementizer *se*:

---

1. At least at the time of writing.

(4) a. *Mi   domando se QUESTO gli   volessero   dire*   (ITA)
          Refl. I-wonder if THIS   to-him(cl) they-wanted to-say
          'I wonder if they wanted to say THIS to him'
     b. **Mi   domando QUESTO se gli   volessero   dire.*
          *Refl. I-wonder THIS   if to-him(cl) they-wanted to-say.

<div align="right">(Rizzi 2001: 289)</div>

Based on the hierarchy in (3), we would expect that the only way for a fronted focus to appear higher than a polarity complementizer would be for the focus to be moved to a left periphery higher than the one in which PolC is merged. In this case, the focus would linearly precede the polarity complementizer simply by virtue of having been dislocated to a different – and crucially higher – CP. This is illustrated in (5), where I use the notation 'CP2' to signal a *higher* left periphery:

(5) [$_{CP2}$ QUESTO mi domando [$_{CP1}$ **se** hanno   detto ~~QUESTO~~]]]
     [$_{CP2}$ THIS   refl. I-wonder [$_{CP1}$ **if** they-have said ~~THIS~~]]]

I will be referring to configurations such as (1) and (4) as relations of *local* precedence (both elements are surfacing in the same CP), and to configurations such as (5) as *non-local* precedence (the two elements are surfacing in different CPs). Based on the hierarchy in (3), we would then expect that, in local configurations, only the α order should be possible, whereas both α and β should be possible in non-local environments.

   In this paper we are interested in determining whether a restriction prohibiting foci from *locally* preceding interrogative complementizers does indeed exist, as argued by Rizzi. To this end, we will be investigating the relative distribution of foci and polarity complementizers in the Slavic language group. We will see that while the β configuration is indeed less frequent than the α order, it is by all means attested. More careful analysis will however reveal that the β order is only possible in those Slavic languages in which the polarity complementizer is a morphologically complex word resulting from the fusion of two independent clausal markers. I will argue that, in these languages, the polarity complementizer is formed through the movement plus incorporation of a lower functional head together with a higher focus marker. Configurations exhibiting the β order are then obtained by giving overt spell-out to the intermediate copy of the movement chain, rather than to its head. In languages where only the α order is possible, on the other hand, the polarity complementizer does not result from the incorporation of two different clausal markers but is rather generated directly in the position in which it ultimately surfaces.

This article is structures as follows. In section II., I show that polarity questions are not islands for focalization.[2] This means that an intervention-effects analysis of the relative order of left-peripheral constituents *à la* Abels (2012) is untenable, for focalization at least (see Callegari 2018). In section III., I investigate word order possibilities in ten different Slavic languages, showing that only three languages in the group (Macedonian, Bulgarian and partially Bosnian-Serbian-Croatian) have β as a grammatical possibility. In section IV., I establish a correlation between languages which display the β order, and languages which have the option of realizing the interrogative polarity particle as an enclitic morpheme attaching to fronted foci. Section V. illustrates the suggested derivation of polarity questions featuring a fronted focus in Macedonian, Bulgarian and Bosnian-Serbian-Croatian; I argue that, in these three languages, the polarity complementizer results from the incorporation of the irrealis marker *da*, which is merged low in the left periphery, together with the focus particle *-li*, which merged as the head of the Focus projection. Section VI. is devoted to discussing the existence of apparently problematic strings of clausal markers; I use these data to further refine my analysis of the irrealis particle *da*. In section VII. I then summarize the contents of the main proposal.

2.   Polarity questions are not Islands for focalization

In his (2012) paper on left-peripheral word order, Abels provides a particularly elegant way of accounting for precedence relations in the left periphery. His logic is the following: no constituent X will ever precede a constituent Y if Y creates an island for X, as the movement of X across Y – necessary to have X precede Y– would result in an intervention effect. We therefore predict that the relative order of these two constituents will always be Y < X. To account for the ungrammaticality of examples like (4b), Abels then argues that a focus cannot precede the polarity complementizer because focalization is blocked by weak islands, and polarity questions, being questions, create weak islands.

This take on word order restrictions introduces complete uniformity between local and non-local precedence relations: if a constituent in narrow focus cannot precede PolC because it is unable to escape the island created by this interrogative element, we expect that configurations where a focus precedes PolC will always be ungrammatical, regardless of whether they are local or non-local. This is because

---

2. I follow Rizzi (1997) in using the term "focalization" to refer to that operation of syntactic dislocation by which a constituent in narrow focus has been fronted to the left periphery of the clause.

movement of the focus across PolC will always result in an intervention effect, regardless of what the landing site for the focus is. Indeed, this is one of the selling points of Abels' analysis: an identical set of principles – intervention effects– can be used to capture both local and non-local word-order configurations.

The idea that focalization is always blocked by the presence of an intervening polarity complementizer however finds immediate counterevidence in the grammaticality of configurations like (6). (6) shows that, in Italian, a focus may non-locally precede *se*, even though the exact same precedence relation is ungrammatical when local, as we saw in (4b).

(6) *QUESTO mi domando **se** gli   volessero   dire   (non qualcos'altro)*
     THIS   refl. I-wonder **if** to-him(cl) they-wanted to-say (not something.else)
     'It's THIS that I wonder whether they wanted to tell him'   (ITA)

If an island violation were behind the ungrammaticality of the local configuration in (4b), we would expect (6) to be equally unacceptable; the fact that (6) is acceptable rather proves that polarity questions are not islands for focalization after all. The existence of this local/non-local asymmetry also shows us that whatever restriction is responsible for (4b) only applies locally, i.e. whenever the focus and PolC surface in the same left periphery.

Note that this local/non-local asymmetry is not peculiar to Romance languages like Italian; it is also found in the Germanic group, as the following examples from Dutch illustrate. Dutch is a V2 language which marginally allows for the fronting of a narrowly focalized constituent to the left periphery of the clause. When fronting occurs inside of a clause containing an embedded polarity question, the only way for the focus to linearly precede PolC is for such precedence relation to be non-local:

(7) a. *Ze vraagt   zich af   **of** MET   JOHN je gesproken hebt (niet MET*
         She wonders refl. vpt. **if** WITH JOHN I spoken   have (not WITH
         MARIE)   (NLD)
         MARY)
         'She wonders whether I have spoken WITH JOHN'
     b. **Ze vraagt zich af MET JOHN **of** je gesproken hebt (niet MET MARIE)*

(8) *MET   JOHN vraagt   ze zich af   **of** je gesproken hebt (niet MET   MARIE)*
     WITH JOHN wonders she refl. vpt. **if** I spoken   have (not WITH MARY)
     'It's WITH JOHN that she wonders whether I have spoken'   (NLD)

3.   Focus and PolC in Slavic Languages

Not all languages pattern with Italian and Dutch in only allowing the β order non-locally: in some cases, a focus may also *locally* precede the polarity complementizer. One such language, as originally noted by Krapova (2002), is Bulgarian, a South-Slavonic language with a pro-drop grammar, no morphological case markers and enclitic definite articles. Consider the following, slightly modified from Krapova's original example:

(10) a. Чудя   се **дали** КНИГИТЕ   Иван ще купи (не   (BUL)
          *Chudja[3] se **dali** KNIGITE   Ivan shte kupi (ne*
          I-wonder refl. **if**   BOOKS.THE Ivan will buy (not
          СПИСАНИЯТА!)
          SPISANIJATA!)
          JOURNALS.THE!)
          'I wonder if Ivan will buy THE BOOKS (not the journals!)'

      b. Чудя   се КНИГИТЕ   **дали** Иван ще купи (не
          *Chudja se KNIGITE   **dali** Ivan shte kupi (ne*
          I-wonder refl. BOOKS.THE **if**   Ivan will buy (not
          СПИСАНИЯТА!)
          SPISANIJATA!)
          JOURNALS.THE!)

Example (10) illustrates how the Bulgarian polarity complementizer *dali* may either precede (10a = α order) or follow (10b = β order) a fronted constituent in narrow focus. Note that the fact that both the α and the β order are acceptable in Bulgarian does not mean that these options are equally unmarked: my informants at least all seem to consistently prefer the order *dali* < FOC, and report that the opposite order feels slightly more archaic.

Example (10) differs from Krapova's original example only with respect to the type of fronted focus involved: while this was simply contrastive in Krapova's example, it is *corrective* (see Bianchi, Bocci & Cruschina 2016) in mine. I chose to use corrective foci because this allows me to draw a full comparison of the various Slavic languages: whereas only some Slavic languages allow for a simply *contrastive* focus to be the target of focalization, most languages allow for the fronting of a *corrective* focus. Indeed, the fronting of a non-corrective focus appears to

---

3. For the transliteration of Bulgarian examples, I am using the Streamlined System (bds 1596: 2009), which is the official transliteration format since 2009 (Selvelli 2015). The one disadvantage of this system is that unambiguous mapping back into Cyrillic is impossible. As I am always providing the original Cyrillic, however, this should not be particularly problematic.

be fairly infrequent in other language groups as well; see for instance Cruschina (2016) on types of focus fronting in Romance languages.

Unsurprisingly, the FOC < *dali* order is also acceptable non-locally. This is shown in (11), where the fronted *KNIGITE* has moved all the way up to the matrix left periphery. Note that the reflexive clitic *ce* has also shifted position, in accordance with second-position requirements on the position of such elements in this language:

(11)　КНИГИТЕ　се　чудя　**дали** Иван ще купu (не СПИСАНИЯТА!)
　　　*KNIGITE　se　chudja　**dali** Ivan shte kupi (ne SPISANIJATA!)*
　　　BOOKS.THE refl. I-wonder **if**　Ivan　will buy　(not JOURNALS.THE!)

Bulgarian is not unique in locally allowing for the order FOC < PolC: this is also grammatical in Macedonian and, for at least some speakers, in Bosnian-Serbian-Croatian (henceforth, *BSC*). I report the relevant examples in (12)–(13). As the reader can see, these are modeled after Krapova's example:

(12)　a.　Не　знам　**дали** КНИГИТЕ　ќе　ги　купи (не　　　　　(MAC)
　　　　　　*Ne[4] znam **dali** KNIGITE　kje　gi　kupi (ne*
　　　　　　Not I-know **if**　BOOKS-THE will them I-buy (not
　　　　　　СПИСАНИЈАТА)
　　　　　　SPISANIjATA)
　　　　　　JOURNALS.THE)
　　　b.　Не　знам　книгите　　**дали** ќе ги　купи (не СПИСАНИЈАТА)
　　　　　　*Ne znam KNIGITE　**dali** kje gi　kupi (ne SPISANIjATA)*
　　　　　　Not I-know BOOKS.THE **if**　will them I-buy (not JOURNALS.THE)

Not all BSC speakers seem to accept the order FOC < PolC locally; I have therefore marked the corresponding example (=13b) with an %.[5]

(13)　a.　*Pitam　se　**da li** će KNJIGE Ivan kupiti (ne ČASOPISE)*　　　(BSC)
　　　　　　I-wonder refl. **if**　will BOOKS Ivan buy　(not JOURNALS)
　　　b.　*%Pitam　se　KNJIGE **da li** će Ivan kupiti (ne ČASOPISE)*
　　　　　　%I-wonder refl. BOOKS **if**　will Ivan buy　(not JOURNALS)

As it was already the case for Bulgarian, both Macedonian speakers and those BSC speakers who find β acceptable seem to prefer the α order over the β one, citing essentially the same reasons provided by Bulgarian speakers: while both options are acceptable, the β option sounds more archaic and stylistically marked than α.

4.　Transliteration of Macedonian has been done according to the ISO9:1995 transliteration standard.

5.　A possible explanation for the overall lower acceptability of the β configuration in BSC will be offered in section V.III.

The β order locally is on the other hand completely ungrammatical in Ukrainian, Polish, Czech, Slovak and Slovene. In these languages, the local fronting of a corrective focus is at least marginally possible -the specific degree of acceptability being dependent on the given language- but only if the focus is internally merged in a position *lower* than the polarity complementizer. Consider for instance (14), from Slovene. In (14), the narrowly focalized object *KNJIGE* may marginally be fronted to the left periphery of the clause, but only if merged lower than the polarity element *če*.

(14) a. *??Sprašujem se,   če bo   KNJIGE Ivan kupil, ne   REVIJE!*   (SLV)
          ??I-wonder refl., **if** will BOOKS Ivan buy,   not JOURNALS!
      b. **Sprašujem se,   KNJIGE če bo   Ivan kupil, ne   REVIJE!*
          *I-wonder refl., BOOKS **if** will Ivan buy,   not JOURNALS!

The examples from Ukrainian, Polish, Czech and Slovak are reported in (15–18):

(15) a. ?Хотів би   я знати,   **чи** КНИЖКИ він купить, не   (UKR)
          *?Xotiv by   ya znaty,   **chy** KNYZHKY vin kupyt',   nie*
          ?Want cond. I to-know, **if** BOOKS   he will-buy, not
          ЖУРНАЛИ!
          ZHURNALY!
          JOURNALS!
      b. *Хотів би   я знати,   КНИЖКИ **чи** він купить, не ЖУРНАЛИ!
          Xotiv by   ya znaty,   KNYZHKY **chy vin kupyt',   nie ZHURNALY!*
          *?Want cond. I to-know, BOOKS   **if** he will-buy, not JOURNALS!

(16) a. *Zastanawiam się,   **czy** KSIĄŻKI Ivan weźmie,   nie MAGAZYNY!* (POL)
          I-wonder   refl., **if** BOOKS Ivan will-take, not JOURNALS!
      b. **Zastanawiam się,   KSIĄŻKI **czy** Ivan weźmie,   nie MAGAZYNY!*
          *I-wonder   refl., BOOKS **if** Ivan will-take, not JOURNALS!

(17) a. *?Zajímá   mě,   **jestli** KNIHY Ivan vezme,   ne ČASOPISY!*   (CZE)
          ?I-wonder refl., **if**   BOOKS Ivan will-take, not JOURNALS!
      b. **Zajímá   mě, KNIHY **jestli** Ivan vezme,   ne ČASOPISY!*
          *I-wonder refl., BOOKS **if**   Ivan will-take, not JOURNALS!

(18) a. *?Zaujíma ma, **či** KNIHY Ivan vezme,   nie ČASOPISY!*   (SLK)
          ?I-wonder refl., **if** BOOKS Ivan will-take, not JOURNALS!
      b. **Zaujíma ma, KNIHY **či** Ivan vezme,   nie ČASOPISY!*
          *I-wonder refl., BOOKS **if** Ivan will-take, not JOURNALS!

We can thus identify two groups in the Slavic subfamily: one the one hand, we have languages like Bulgarian, Macedonian and partially BSC, in which the β

order is a grammatical possibility even locally. Note that these all feature an identical – spelling differences aside– polarity complementizer: this is *dali* in MAC and BUL, and *da li* in BSC.

On the other hand, we have languages like Ukrainian, Polish, Czech, Slovak and Slovene, where a focus cannot locally precede the polarity morpheme. Note that even in these languages we observe the same local/non-local asymmetry observed for Dutch and Italian; the example in (19), from Slovene, shows how focus fronting across PolC and to a higher left periphery is possible.

(19) *(Ne, ne,) KNJIGE se   sprašujem **če** bo   Ivan kupil, (ne*   (SLV)
     (No, no,) BOOKS refl. I-wonder **if** aux.fut Ivan buy,   (not
     REVIJE)!
     JOURNALS)!

Ukrainian, Polish, Czech, Slovene and Slovak thus behave exactly like Dutch and Italian: the β order is possible only if the focus has been fronted to a left periphery different from the one in which PolC surfaces.

I will be referring to the first language group (languages which have β as a grammatical possibility locally, namely Bulgarian, Macedonian and BSC) as *group I*, and to the latter group (locally, only α is possible) as *group II*. Group II languages behave exactly like expected given the hierarchy in (3), whereas the existence of group I languages is unexpected.

A potential solution which might be put forward to account for the grammaticality of the β pattern locally (group I languages) would be to argue that those nominal expressions appearing in a pre-*dali/da li* position ((b) examples in (10) through (13)) are in fact topics and not foci. What is problematic about such analysis is that (10)–(13) clearly depict a corrective speech act, and corrections are arguably always focal, not topical.[6] Still, this is an argument worth spending some

---

6. Regardless of what definition of *focus* one picks, the notion of "topical correction" makes little sense. Consider the exchange below, where B' contribution to the communicative exchange is clearly corrective in nature:

(i)   A: Mary likes Ricardo
      B: Mary likes KAREN (not RICARDO)!

Under a definition of "focus" as the part of the utterance which represents new information (Halliday 1967; Chomsky 1970; Jackendoff 1972), it is clearly the correction itself which qualifies as such: everything else in B's utterance is a repetition of what was originally stated by A. If we take *focus* to correspond to the most informative part of the sentence (Roberts 1996), once again it is the direct object *KAREN* which qualifies as such, if anything because this is the only constituent representing new information. Concerning semantic definitions of *focus, KAREN* is also the only linguistic expression in (i) whose interpretation gives rise to a set of alternative propositions (Rooth 1992, 1999).

time on, as it would save us from the predicament we currently find ourselves in: according to the hierarchy in (3), topics *can* precede polarity complementizers. In particular, we see that in (3) there is at least one Topic projection higher than IntP.

A topic analysis of pre-*dali* constituents would seem to find support in the fact that clitic doubling of the fronted *KNIGITE* is mandatory in the Macedonian example in (12). In several languages, clitic resumption is intimately connected with topicality; such is the case in Romance languages, where constituents which represent old and presupposed information generally feature mandatory clitic resumption (see Cruschina 2010 on clitic resumption possibilities in Romance). The presence versus lack of a coindexed clitic is however regulated by mechanisms other than Information Structure in Macedonian: in this language, the relevant criterion is definiteness. As a matter of fact, in Macedonian, indirect and direct objects which are definite DPs, proper names or NPs modified by demonstratives are obligatorily clitic-doubled. Indefinite DPs in object positions on the other hand cannot be clitic-doubled (Rudin et al. 1999; Kochovska 2010). This is illustrated in (20): clitic-doubling of the object is mandatory in (20a), but ungrammatical in (20b):[7]

(20) a. Иван *(ja) прочитал книгата / оваа книга   (MAC)
          *Ivan *(ja) procital   knigata   / ovaa kniga*
          Ivan   *(it) read   book-the / that book
          'Ivan read the book/that book'

                                                        (Kochovksa 2011: 242)

      b. Иван *(ja) бара   една секретарка
          *Ivan *(ja) bara   edna sekretarka*
          Ivan   *(her) look-for one   secretary
          'Ivan is looking for a secretary'

                                                        (Kochovksa 2011: 245)

Something we can do to investigate whether pre-*dali/da li* could be topics is determining whether non-referential expressions such as 'enough X' or 'at least an X', which are standardly considered never to be topical, can be fronted to a position preceding *dali/da li.* It turns out they can, in all three languages. Starting with Macedonian (21), we see that an expression like 'enough books' may appear both before and after *dali/da li.* Note that this constituent is not clitic-doubled by a cor-

---

7. A similar process seems to be taking place in Bulgarian: in colloquial Bulgarian at least, [+definite] DPs such as proper names and nominal expressions featuring the definite enclitic article generally appear together with a resumptive clitic element agreeing in person, number, case and gender (Vakareliyska 1994; Krapova 2002). Unlike Macedonian, where clitic resumption of definite nominal expressions is mandatory, this process is still optional in Bulgarian.

responding clitic like *gi* (=them), precisely because this is not a definite expression:

(21) a. Се   прашувам **дали** ДОВОЛНО КНИГИ ќе   купи, (не   (MAC)
         *Se   prašuvam **dali** DOVOLNO KNIGI   ḱe kupi (ne*
         Refl. I-wonder **if**   ENOUGH   BOOKS will buy (not
         ДОВОЛНО СПИСАНИ**ЈА**!)
         DOVOLNO SPISANI ĵA!)
         ENOUGH   JOURNALS!)
         'I wonder if s/he will buy ENOUGH BOOKS (not ENOUGH JOURNALS!)'
     b. Се   прашувам ДОВОЛНО КНИГИ **дали** ќе   купи, (не ДОВОЛНО
         *Se   prašuvam DOVOLNO KNIGI **dali** ḱe kupi, (ne DOVOLNO*
         Refl. I-wonder ENOUGH   BOOKS **if** will buy (not ENOUGH
         СПИСАНИЈА!)
         SPISANIJA!)
         JOURNALS!)

Examples (22)–(23) detail how the same holds for BSC and Bulgarian: expressions such as 'at least a book' and 'enough books' may appear both before and after *dali/da li.*

(22) a. *Pitam   se **da li** je BAR   KNJIGU Ivan kupio,   ne*   (BSC)
         I-wonder refl **da li** aux AT.LEAST BOOK   Ivan bought, not
         BAR   ČASOPIS
         AT.LEAST JOURNAL
         'I wonder if he bought AT LEAST A BOOK, not AT LEAST A JOURNAL!'
     b. *Pitam   se BAR   KNJIGU **da li** je Ivan kupio, ne BAR*
         I-wonder refl. AT.LEAST BOOK   **da li** aux Ivan bought, not AT.LEAST
         ČASOPIS
         JOURNAL

(23) a. Чудя   се **дали** ДОСТАТЪЧНО КНИГИ той ще kynu, (не (BUL)
         *Chudya se **dali** DOSTATACHNO KNIGI   toy shte kupi, (ne*
         I-wonder refl. **if** ENOUGH   BOOKS he   will buy, (not
         ДОСТАТЪЧНО   СПИСАНИЯ!)
         DOSTATACHNO SPISANIYA!)
         ENOUGH   JOURNALS!)
         'I wonder if she will buy ENOUGH BOOKS (not ENOUGH JOURNALS!)'

b. чудя   се ДОСТАТЪЧНО КНИГИ **дали** той ще kynu, (не
*Chudya se DOSTATACHNO KNIGI **dali** toy shte kupi, (ne*
I-wonder refl. ENOUGH   BOOKS **if**   she will buy, (not
ДОСТАТЪЧНО СПИСАНИЯ!)
DOSTATACHNO SPISANIYA!)
ENOUGH   JOURNALS!)

We can conclude that constituents which are fronted to a pre-*dali/da li* position can indeed be foci in all languages which make up group I.

3.1 Overview

Table 1 provides an overview of the Slavic languages reviewed so far, offering a bird's-eye view of which languages locally allow for the β order,[8] and which ones do not. Languages are divided according to their respective subgroup:

Table 1. Beta order in Slavic languages

Subgroup	Language	Allows for β?
South Slavic	BSC	%YES
South Slavic	Bulgarian	YES
South Slavic	Macedonian	YES
South Slavic	Slovene	NO
West Slavic	Czech	NO
West Slavic	Slovak	NO
West Slavic	Polish	NO
East Slavic	Ukrainian	NO

Based on Table 1, we can conclude that the possibility of having a constituent in focus locally precede the polarity complementizer is the exception rather than the norm, with 5 out of 8 languages not having this configuration as a grammatical possibility. Table 1 also shows that the β order is geographically circumscribed to the South Slavic area, although it is by no means a feature of all the languages in this particular subgroup: Slovene rather patterns with Czech, Slovak, Polish and Ukrainian in only displaying the α order.

---

8. Unless otherwise stated, from now on all mentions to the α and the β orders are to be understood as *local* instances of these configurations.

One language notably missing from Table 1 is Russian. This is because Russian has no free morpheme to realize the polarity complementizer, but rather employs the enclitic particle *-ли* (=*-li*). Russian *-li* must encliticize onto a finite verb (24) or a fronted narrow focus if a focus is present (25):

(24) Я не знаю, работает **ли** он на заводе   (RUS)
      *Ja ne znaju, rabotaet **li** on na zavode*
      I   not know, works   -**li** he at factory.PREP
      'I don't know if he works at the factory'

(25) КНИГУ   **ли** он читает?[9]   (RUS)
      *KNIGU   **li** on čitaet?*
      BOOK.ACC -**li** he reads?
      'Is it THE BOOK that he is reading?'

I will be assuming that both polarity complementizers and *-li* realize the category POL; a POL element signals that a given clause is to be interpreted as a polarity question. In Russian, the relative order of fronted narrow foci and the morpheme expressing POL is then always fixed in the β configuration as a result of the enclitic nature of POL itself.

Another East Slavic language missing from Table 1 is Belarusian. Belarusian does not feature in this table for the same reason Russian does not: this language has no free morpheme to express POL. Exactly like Russian *-li*, Belarusian ці- (= *ci-*) takes as its prosodic host either the fronted finite verb, or a constituent in narrow focus. Unlike *-li*, however, *ci-* is *pro*clitic on its phonological host:

(26) Цікава,   **ці** чытала яна кнігу   (BEL)
      *Cikava, **ci** chytala jana knigu*
      I-wonder, **ci**- read   she book.ACC
      'I wonder if she read the book'

(27) Цікава,   **ці** КНІГУ   яна чытала   (BEL)
      *Cikava, **ci** KNIGU   jana chytala*
      I-wonder, **ci**- BOOK.ACC she read
      'I wonder if it is THE BOOK that she read'

With Belarusian, we thus observe a situation opposite of that observed for Russian: the relative order of the morpheme expressing POL and fronted foci is fixed in the α configuration as a result of the proclitic nature of *ci-*.

---

9. In matrix environments, the most pragmatically neutral way of forming a polarity question in present-day Russian is through intonation alone. The *-li* particle is on the other hand obligatory in embedded yes/no questions.

4.   Enclitics & complex complementizers

In the previous section, we saw that Slavic languages can be divided in two groups depending on whether a fronted focus may or may not locally precede a polarity complementizer. On the one hand, we have languages like Macedonian, Bulgarian and, at least for some speakers, BSC (group I languages), which have this as a grammatical option. On the other hand, we have languages like Slovene, Czech, Slovak, Polish and Ukrainian, where, exactly like in Italian and Dutch, a locally fronted focus may only *follow* PolC. What is responsible for this specific divide?

A first factor differentiating the two groups is the possibility, attested in all group I languages, to also realize POL as the enclitic morpheme *-li*. Recall that, when lexically realized, POL in Russian takes the form of the enclitic morpheme *-li*, attaching either to the fronted verb or to a fronted focus. Macedonian, Bulgarian and BSC have both options: on the one hand, they can realize POL as a free particle, exactly like in Italian and Dutch. In this case, as we already saw, POL takes the form of *dali*/*da li*. On the other hand, POL can also take the form of the enclitic morpheme *-li* (a cognate of Russian *-li*), attaching either to the finite verb or to a constituent in narrow focus, exactly like its Russian counterpart. That POL may also be realized as an enclitic morpheme is perhaps not surprising, as several elements are realized as clitics in Slavic languages: the inventory consists not of simply pronominal objects, but also of auxiliaries of tense and aspect, negation and, as we have just seen, question particles.

As a rule, the enclitic *-li* polarity strategy may be employed to form both matrix and embedded yes/no questions. The use of the *-li* strategy is illustrated in (28)–(29) for matrix polarity questions, and in (30) for embedded ones:

(28)   ИВАН ли рисува всеки ден?   (BUL)
        IVAN li risuva vseki den?
        IVAN -li paints every day?
        'Is it IVAN who paints every day?'

                                                            (Dukova-Zheleva 2010: 177)

(29)   Рисува ли Иван всеки ден?   (BUL)
        Risuva li Ivan vseki den?
        Draws -li Ivan every day?
        'Does Ivan draw every day?'

                                                            (Dukova-Zheleva 2010: 170)

(30)   *Pitam   se   studira li Marko medicinu*   (BSC)
        Wonder-I refl. studies -li Marko medicine
        'I wonder if Marko studies medicine'

Likewise, *dali/da li* may be employed to form matrix as well as embedded polarity questions.

Both *dali* and *li* thus signal that a given proposition is to be interpreted as a polarity question; in this sense, the two elements are functionally equivalent. This does not mean that the two elements are also *pragmatically* equivalent: in both Macedonian and Bulgarian, whereas *dali* is reported to be the most pragmatically neutral option to form an embedded interrogative, embedded *-li* questions are perceived to convey something akin to direct quotation or free indirect speech.[10] Differently put, whereas an embedded *dali* question is closer to the run-of-the-mill case of embedded interrogative, embedded *-li* questions come closer to matrix interrogatives in exhibiting something akin to their own illocutionary force.[11]

The existence of a second, enclitic strategy to realize polarity questions, and the fact that this is attested in all group I languages, leads us to formulating the following generalization:

(31)   Only those Slavic languages which have the option of realizing POL as the enclitic morpheme *-li* have the option of fronting a constituent in narrow focus to a position locally preceding the polarity complementizer.

Why would a generalization (31) be relevant to capture the divide between group I and group II languages? The proposal will be fleshed out in more detail in the following section, once we have had a chance of appreciating all the necessary data. For now, let us simply consider the basic intuition: assuming that complementizers like *dali* and enclitic morphemes like *-li* are both instantiations of POL, those

---

10. Note the interesting parallel between embedded *-li* questions in Slavic, and embedded V2 in those Germanic languages that only have mandatory V2 in matrix contexts -such as Norwegian-. In the latter languages, V2 in an embedded declarative clause appears to signal independent illocutionary force, something which is reminiscent of the free indirect speech "feel" we observe in Slavic embedded *-li* questions.

11. The relative frequency of the *-li* strategy as opposed to the *dali/da li* one (or even to the lack of any polarity marker in languages where this is possible, as it is the case in Macedonian) also depends on the specific language. Variation is especially observed in matrix questions. Take the otherwise two closely related languages Macedonian and Bulgarian: in Bulgarian, *-li* questions are the unmarked option to form a matrix polarity question, whereas in Macedonian, the preferred strategy to form a matrix polarity question is through intonation alone. In Englund's (1977) Macedonian corpus study, for instance, only 30% of all polarity questions are *-li* questions, whereas 44% has no morphologically realized (neither *-li* nor *dali*) polarity marker at all (see also Friedman 1993, and Rudin et al. 1999). This clashes with the results provided by Nikov (1976) for Bulgarian, who found that 93% of all instances of matrix non-rhetorical polarity questions featured *-li*.

languages which allow for a PolC to follow a fronted focus are the same languages which already allow for (a different kind of) POL to follow a constituent in focus.

In its current formulation, (31) is too permissive: Czech also has the option of realizing POL as the enclitic *-li* morpheme, and yet this is one of the languages in which only the α order is possible. The enclitic *-li* polarity strategy in Czech is illustrated in (32), which I take from Toman (1996):

(32) *Nevíme,   mají   li dnes medovinu*   (CZ)
       Not-we-know, they-have -li today honey
       'We don't know whether they have honey today'

<div align="right">Toman (1996: 508)</div>

Clearly, some refinement to (31) is needed. I contend that the additional factor playing a role in licensing the β order is the morphological composition of PolC itself, specifically whether this is morphologically transparent.

As already briefly remarked in section III., all languages which allow for β display an almost identical PolC: this is *dali* in MAC and BUL, and *da li* in BSC. Historically, this element has been analyzed as resulting from the union of the irrealis morpheme *da* together with the question particle *-li* (Hansen et al 2016). Note that both *da* and *li* are still used in isolation and are fully functional clausal markers in present-day MAC, BSC and BUL. Precisely because these two morphemes are still fully functional even in isolation, I want to extend Hansen et al.'s diachronic-incorporation analysis of *dali/da li* to the synchronic dimension, and argue that present-day *dali/da li* results from the incorporation through movement of the irrealis clausal marker *da* together with the enclitic polarity morpheme *-li*. As we will see in section V, it is precisely this process of incorporation which gives PolC a landing site lower than the Focus projection.

What about Czech *jestli*? Unlike Macedonian, Bulgarian and BSC *dali/da li*, Czech *jestli* does not arise from the union of two subordination markers that are still functional in present-day language, but rather from the union of *-li* together with the verbal form *jest* (Tabakowska 1997). Crucially, *jest* corresponds to the *former* 3rd person singular of the verb to be (*být*), and hence is a form no longer use in present-day Czech.

We can then revise the generalization in (31) as in (33):

(33) Only those Slavic languages (i) which have the option of realizing POL as the enclitic morpheme *–li, and* (ii) where PolC results from the union of markers which are still functional even in isolation, have the option of having a fronted focus locally precede PolC.

5.   *Da* Meets *–li,* polarity ensues

Before we discuss the proposed derivation for *dali/da li* structures, let us analyze the properties and distribution of the clausal marker *da*.

The proclitic morpheme *da* is found in all South-Slavic languages and is an incredibly flexible element which either participates or single-handedly results in the expression of a remarkably wide range of different constructions. These range from the marking of non-factive subordination to the formation of negated future tenses, the expression of counterfactual clauses, the expression of deontic modality and epistemicity, and the expressions of wishes, desires or commands (optatives). Below, we review some of these functions.

A first function performed by *da* relates to the expression of non-factivity. Slavic languages morphologically mark the distinction between subordinating conjunctions introducing factive embedded clauses and those introducing non-factive ones. Unlike what happens in languages like Russian, where the non-factive complementizer is a morphological variant of the factive complementizer,[12] the non-factive marker in BSC, MAC and BUL is a separate morpheme. In all these three languages, this is *da.* This morpheme is in overt opposition and hence complementary distribution with complementizers introducing factive clauses; these are *deka* in Macedonian, *če* in Bulgarian, and *što* in BSC.

In Macedonian and Bulgarian, in those environments where both *da* and *če/ deka* are possible, we see that while *če/deka* marks an independently true statement, *da* carries no such implication. This is exemplified in (34) for Bulgarian, which I take from Rudin (1986). Note in particular that *če* must be used if the speaker was indeed hungry:

(34) a. Не усещах че съм гладен   (BUL)
          Ne useshtah che sum gladen
          Not I-noticed that I-am hungry
          'I did not notice that I was hungry (even though I was)'

      b. Не усещах да съм гладен
          Ne useshtah da sum gladen
          Not I-noticed that I-am hungry
          'I did not notice that I was hungry (because I probably wasn't)'

                                                              Rudin (1986: 58)

---

12. i.e., *чтобы (čtoby),* which is obtained by incorporating the declarative complementizer *čto* together with the conditional marker *by*.

BSC *što* is slightly more specialized than Bulgarian *če* and Macedonian *deka,* but differences in the use of *da* as opposed to *što* can still be reduced to differences in perceived factivity. In particular, whereas *što* introduces the clausal complement of emotive factive verbs (as well as relative clauses), *da* is used for all other types of clausal complements:

(35) a. *Ana misli   da   Marko spava*   (BSC)
       Ana thinks that Marko sleeps
       'Ana thinks that Marko is sleeping'
     b. *Ani   smeta   što Marko stalno spava*
       Ana.dat bothers that Marko always sleeps
       'It bothers Ana that Marko is always sleeping'

<div align="right">(Mihaliček 2012: 114)</div>

Note that neither Macedonian nor Bulgarian possess a system of infinitives like English does;[13] in these two languages, what would be an infinitive structure in English is realized by having *da* procliticize onto the embedded verb. This is illustrated for Macedonian with the contrast between (36a) and (36b). Note that the verb onto which *da* procliticizes still retains person and number features:

(36) a. Заборавив дека седам во канцеларија   (MAC)
       Zaboraviv   deka sedam vo kancelariǰa
       I-forgot   that I-sit   in office
       'I forgot that I was sitting in an office.'
     b. Прозорецот заборавив да   го затворам
       Prozorecot   zaboraviv   da go zatvoram
       The-window I-forgot   that it I-close
       'I forgot to close the window.'

<div align="right">(Lindstedt 2010: 416)</div>

Following much existing literature on *da* in South Slavic languages, and for lack of a better term, I will be referring to structure like that in (36b) -where *da* introduces as embedded clause- as *subjunctive da* structures.[14]

A second function relates to the expression of modality. When *da* appears sentence-initially, it signals the presence of a desire or a weak command (37a). In

---

13. The same also holds for at least some varieties of BSC: Torlakian dialects (south-eastern part of Serbia) for instance generally pattern with MAC and BUL in only having the *da* option to form an "infinitive" structure. See Joseph (1983) for an excellent overview of the loss of infinitives in Bulgarian and Macedonian, and the partial loss of the same structure in BSC.

14. See however Lindstedt (2010) on why using the term "subjunctive" may be problematic.

this case, *da* performs a function similar to that performed by subjunctive mood in matrix sentences in languages like Italian (37b):

(37) a. Да живее Европейскиям съюз   (BUL)
          Da zhivee Evropejskiyat   sajuz
          Da lives European.the union
          'Long live the European Union!'

      b. *Che tu possa   vivere a lungo!*   (ITA)
          That you may(subv) to-live a long
          'May you live a long life!'

In matrix polarity questions, on the other hand, the presence of *da* correlates with a deontic interpretation of the event whose polarity one is questioning (38):

(38) Да се върне   ли довечера?   (BUL)
      Da se   varne   li dovechera?
      Da refl. (s)he-returns li tonight?
      'Should (s)he return tonight?'

                                                                  (Rudin 1986: 118)

Finally, *da* may either be used other on its own or in combination with the subjunctive marker *bi* to introduce a counterfactual clause:

(39) Да го добиевме   твоето писмо, ќе дојдев   (MAC)
      Da go dobievme   tvoeto pismo, ќe dojdev
      Da it we-had-received your   letter,   will I-had-come
      'If we had received your letter, I would have come'

Summing up, *da* appears to perform a multitude of different functions in the three languages under discussions, functions which we may characterize more in general as pertaining to the expression of [−realis] environments. In particular, *da* performs a function similar to the one performed by the infinitive English marker *to* in embedded domains, and to what Romance languages express through the use of subjunctive mood in matrix propositions.

What about the distribution of *da*? At least in MAC and BUL, *da* must follow any fronted focus or topic, as shown in (40). In (40a), the embedded subject occurs in a post-verbal position; if locally fronted through focalization or topicalization, this must crucially precede *da* (40b) vs. (40c).

(40) a. Искам да пеят децата   (BUL)
          Iskam da peyat detsata
          I-want da sing   children-the
          'I want the children to sing'

      b. *Iskam detsata da peyat*

    c.   *Iskam da detsata peyat

<div align="right">(Rudin 1983: 4)</div>

In fact, precisely because of the proclitic nature of *da*, nothing other than other clitics may intervene between *da* and the verb onto which this procliticizes. This includes the pronominal subject of the embedded clause, which must either precede *da* (41a) or follow the embedded verb (41c).

(41) a.   Искам мя да дойде   (BUL)
          Iskam tya da doyde
          I-want she da comes
          'I want her to come'

<div align="right">(Lindstedt 2010: 416)</div>

    b.   *Iskam da tya doyde*
    c.   *Iskam da doyde tya*

Note that *da* is in complementary distribution with other markers of clausal subordination, with the single exception of relative pronouns in Bulgarian at least. This has led Lindstedt (2010) to argue that *da* has full complementizer status.[15]

5.1   The derivation

(42) illustrates the proposed derivation for sentences featuring *dali* and a fronted narrow focus. (42) is meant to illustrate the process resulting in the formation of an embedded polarity question in three languages (MAC, BUL and BSC) which, despite the surface similarities, are still quite different; as such, (42) is inevitably underspecified with respect to some features.[16] The derivation in (42) is based on a cartographic (Rizzi 2001 in particular) understanding of the structure of the left periphery; this is because Rizzi's template provides a particularly neat way of illustrating what I believe to be the mechanism underlying the formation of PolC in the languages under discussion. Note however that nothing in my analysis hinges on assuming that specific types of constituents are associated with dedicated functional projections in the left periphery, which is one of the main tenets of cartography. Where my analysis does rely on Rizzi's hierarchy, on the other hand, is in

---

15. See Shlonsky's (1988) paper on Hebrew's *še* for another example of complementizers which are phonetic clitics.

16. Some features that the derivation in (43) does not immediately capture are for example:

–   The fact that in BSC, *-li* is a strict second-position clitic, whereas in Bulgarian and Macedonian, *-li* may surface in positions other than the classic Wackernagial second position.

–   The fact that *-li* in Macedonian may split a focused DP, whereas in Bulgarian *-li* follows the entire DP even in those cases where only part of such constituent is focalized.

assuming that the ultimate landing site of PolC will be higher than the position in which foci are internally merged; other than the order FOC < PolC being cross-linguistically rarer than PolC < FOC, this is something for which I have not yet found independent justification.[17]

In (42), *da* is externally merged in the head of Fin. *Da* then moves up, combining with the enclitic particle -*li*, which I take to be merged as the head of the Focus projection -as we will see in V.II-. The resulting complex form *dali/da li* is then moved to the head of IntP, the position in which polarity complementizers are hosted according to Rizzi (2001):

(42)
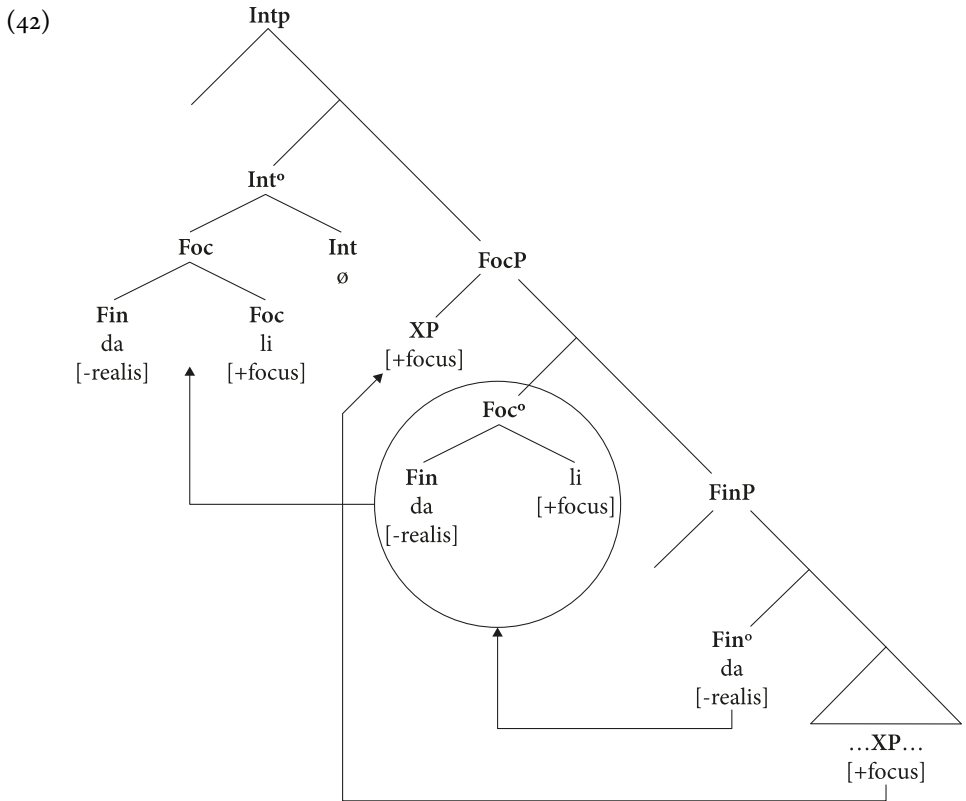
Let us discuss the details of (42), starting with the proposed external merge site for *da*.

In (40)–(41), we saw that, in Bulgarian and Macedonian, *da* must follow embedded topics as well as the embedded subject. We also know that, when occurring together with a relative pronoun (a possibility in BUL at least), *da* must follow said relative pronoun. If we are to account for the low position of *da* while at the same time maintaining that this element has complementizer status, as argued by Lindstedt (2010), a sensible option would be to take it to originate in *Fin*, the lowest functional projection capable of hosting complementizers in Rizzi's hierarchy (see again (3)). This seems sensible also because, in Rizzi's framework, Fin is the functional projection dedicated to hosting [−realis] subordination markers.

As already mentioned in the preceding section, South-Slavic *da* is not a free morpheme but a clitic. Unlike *–li, da* is *pro*clitic: it must linearly *precede* its phonological host. In most cases, what provides *da* with a site onto which to procliticize is a verbal head: this will result in the verb-adjacency pattern we observed in (41). Verbs are not the only elements which may provide *da* with a phonological host: in some cases, a second clitic may also do. I argue that this is exactly what happens in *dali/da li* structures: *da* moves up from the head of Fin to combine with the enclitic *-li* morpheme, giving rise to the complex morphological word *dali/da li*, and providing a phonological host for both *-li* and *da* in the process.[18, 19]

---

18. Following Richards (2010, 2016), I am assuming that at least some phonological features are visible in the syntax and that some kinds of syntactic displacement operations are motivated by the need to derive a well-formed prosodic structure (see also Samek-Lodovici (2015) and Szendrői (2002, 2003, 2017) on left dislocation of focalized constituents being triggered by prosodic requirements). In the case at hand, I am assuming that the information pertaining to the enclitic/proclitic nature of the clitic itself is visible to the narrow syntax and hence can trigger syntactic movement.

19. (43) illustrates the derivation of *dali* questions, which are the focus of this paper, but do not forget we also have questions of the type of (29)–(30) (which I repeat below), where the element immediately preceding -li–i.e., the element onto which *-li* encliticizes–is either the fronted verb (i) or a fronted focalized expression:

(i) **Risuva** li Ivan vseki den?
     Draws -li Ivan every day?

(ii) **IVAN** li risuva vseki den?
      IVAN -li paints every day?

Concerning (i), I am assuming that the verb moves to Foc°, essentially replacing *da* in (43). In this sense, both *da* and the fronted verb perform an identical function: they provide a phonological host for the enclitic *-li* morpheme. Concerning (ii), I take the DP to front to the specifier of the Focus projection, exactly like depicted in (43) for the fronted focus in structures where

The process of double-cliticization (i.e., *-li* encliticizing onto *da*, and *da* procliticizing onto *-li*) I am suggesting to be underlying *dali/da li* is by no means unheard of in Slavic languages. Consider in particular the negation morpheme *ne*. In both Bulgarian and Macedonian, *ne* is proclitic onto the verb: nothing with the exception of other clitics[20] may appear in between *ne* and the verb this morpheme procliticizes onto. As (43) shows, in MAC *-li* can encliticize directly onto the negation itself; this results in a negative-polarity-question interpretation.

(43) Не ли сакаш   да одиш?   (MAC)
      Ne li sakaš   da odiš?
      Not -li you-want da you-go
      'Don't you want to go?'

(Rudin et al. 1999: 556)

(43) thus shows that a second clitic can indeed function as a host for *-li*.

Recall that, in BUL and MAC at least, *dali* looks like a full-fledged word and not simply as the juxtaposition of two distinct morphemes. Could two separate clitics give rise to a single word? They can, and once again we need not look further than South Slavic languages themselves to find evidence for this kind of process.

An anonymous reviewer wonders whether the fact that BSC *da li* is spelled as two separate morphemes should be taken as evidence that a different derivation underlies the PolC in BSC. While this is certainly a relevant question, one should also note that orthographic conventions are not perfect indicators of the word or non-word status of a given linguistic expression. Cross-linguistically, we know there is considerable variation wrt the orthography of clusters of clitics, with speakers often being unsure about how these should be spelled. Consider the case of Italian *glielo*, resulting from the combination of the dative masculine pronominal element *gli* with the masculine accusative pronoun *lo*. Cardinaletti (2008) argues that *glielo* results from the adjunction of *gli* and *lo* onto one another under an identical functional head. The author points out that children (as well as a good number of adults, I might add) often spell it as *glie lo* (or variations thereof), and that the Italian writer Italo Calvino (1923–1985) spelled it the very same way.[21]

---

da is present. The difference between (ii) and (44), then, only relates to the presence of *da*: the focused DP always fronts to the same projection.

20.   The exact relative order of clitics inside the clitic cluster is the following:

(i)   Negation > Auxiliary > Dative > Accusative > VERB

21.   Note that a similar debate surrounds *ne li*. According to Tomić (1996), for instance*, ne li* written as two separate morphemes (as in (44)) is actually an error: only *neli*, written as a single word, exists.

Cardinaletti suggests taking orthographic conventions to go one way only: if two elements are written as a single word, they are one constituent. If they are written separately, they may be one or multiple constituents. This take on orthographic conventions also accounts for the fact that the clitic cluster *me lo* (1st person dative + 3rd person accusative) behaves identically to *glielo* wrt both syntax and prosody, as Cardinaletti shows; this is even though the two expressions are spelled differently. It could very well be, then, that BSC *da li* is interpreted a word even though it is not spelled like one. This analysis is not so far-fetched given that it has in fact been suggested before (see Radanović-Kocić 1988 and Hock 1992) and references therein for arguments and counterarguments).

We have now accounted for why *da* would move up to combine with *-li*: by virtue of this process, both the proclitic *da*, and the enclitic *-li* receive a phonological host. The derivation is however not complete: if the morphologically complex word *da li/dali* were to remain in the position in which its two morphemes are combined (i.e., the head of FocP), we would expect PolC to always follow constituents in focus in BSC, MAC and BUL. Differently put, we would expect the α order never to be possible, a hypothesis which is falsified not only by the fact that MAC, BUL and BSC all display this configuration, but also given the fact that the α order seems to be the least marked option. To account for this intra-linguistic preference, as well as for the fact that, cross-linguistically, polarity complementizers tend to precede constituents in narrow focus, I thus take the complex *da li/dali* to move further up in the left periphery. A candidate position for the landing site of this additional movement step is the head of IntP, which we saw to be the functional projection specialized for hosting PolC according to Rizzi (2001). We have now derived the order PolC < FOC, and the derivation is now complete.

5.1.1   *Where does -li go?*

In (42), *-li* is the head of the Foc(us) projection.[22] Assuming that *-li* is connected to the expression of focus seems sensible given the existence of a clear link between *-li* and constituents in focus. We already saw that, if a constituent in narrow focus is present, this must front to a position immediately preceding *-li*. Conversely, it is always possible to interpret pre- *-li* constituents as being in focus, regardless of whether they are verbal or nominal in nature. Consider (44):

---

22. An identical line of analysis is adopted by Dukova-Zheleva (2010). See also Rivero (1993), who suggests *-li* to be base-generated directly in C°.

(44) Гледа ли го?   (BUL)
      Gleda li go?
      Looks -li it(cl)?
      'Is (s)he looking at it?'

-Li questions in which what is fronted is the finite verb have traditionally been referred to as "neutral", in a bid to distinguish them from DP-li questions, which result in a narrow-focus interpretation of the fronted DP. In fact, *yes/no* questions with the structure of (45) can be equally – context permitting– interpreted as being narrowly focalized: a possible interpretation of (44) is that according to which the verb is explicitly contrasted to a relevant set of alternative activities. This is possible precisely because the verb fronts to a position immediately preceding – li:

(45) Is she *looking* at it? (as opposed to some other relevant action *x*)

Note also that in Russian, inherently non-focalized indefinites such as *kto-nibud'* somebody/anybody) and *čto-nibud'* (something/anything) may never be fronted to a pre- *-li* position, thus providing additional evidence of the close connection of *-li* with focalized expressions.

　　If *-li* is indeed generated in the head of FocP, we can conclude that this morpheme minimally spells out the presence of a [+ focus] feature. Additionally, one could also assume that *-li* also has a [+ interrogative] feature: after all, we have seen that the presence of *-li* signals the clause to be a polarity *question*. This latter line of analysis, which has been adopted by King (2001) and Dukova-Zheleva (2010) among others, would be corroborated by the lack of any additional interrogative morpheme specifically marking the clause as being a question. Evidence against a superimposing of the [+focus] and [+interrogative] features when analyzing *-li* comes from the following Macedonian example, which I take from Friedman (2018). (46) illustrates how *-li* may also have a purely emphatic, *non*-interrogative interpretation: in (46), *-li* is used in conjunction with the expression of a dubitative structure.

(46) *Kako da ne, toj **li** kje ti   ja napravi kolata...*   (MAC)
      How da not, he **li** fut to.you it fixes   car.the...
      'Oh sure, *he*'ll fix your car alright...'

                                                          (Friedman 2018: 46)

According to Friedman (p.c.), while (46) is considered a bit of a "Serbism" stylistically speaking, it is perfectly acceptable in Macedonian.

　　An alternative line of analysis would be to assume that *-li* is specialized for [focus] alone, and that what is responsible for the typing of the structure as a polarity question is a null [+Q] morpheme in a distinct projection in the CP (e.g.,

Force). In this respect, consider the case of Jula, which equally possesses a dedicated morpheme to mark the presence of narrow focus. This morpheme, *lo*, is crucially very different from the morpheme used to mark a proposition as being a polarity question, *wa* (see [Callegari 2021]). This is illustrated in (47a)–(b). Interestingly, *wa* (unlike *lo*) need not be given overt spell-out to type the clause as being a yes/no question: very often, intonation alone will do.

(47)   a.   *Seydou ka   dji   mi   (wa)?*   (JUL)[23]
            Seydou IMPF water drinks (ptc)?
            'Is Seydou drinking water?'
       b.   *Seydou lo ka   dji   mi*
            Seydou ptc IMPF water drinks
            'It is Seydou who drinks water'

In this paper I will then simply assume that *-li* minimally spells out a [+focus] feature; I leave open the question of how to reconcile that with the fact that the vast majority of *-li* environments are interrogative environments.

   If *-li* is specified as being [+focus], and *da-* marks the presence of a [−realis] environment, the resulting *dali/da li* complementizer results as being specified for both features. Intuitively, this seems like a sensible featural composition for an element whose function is that of introducing polarity questions: being interrogative in nature, *dali* structures involve the computation of a set of alternatives answers, something which has been analyzed in terms of focal alternatives (see Rooth 1992, 1999). Since polarity questions obviously do not presuppose the veridicality of their propositional content, they are also [−realis] environments *par excellence*.

5.2   Optionality

How does (42) capture the fact that group I languages may display both the α *and* the β order?

   Consider again the tree in (42). There is a point in the derivation in which the complex morpheme *dali* is found in a position *preceding* the position in which the focus is internally merged: this occurs in the head of FocP, namely in the projection where the two morphemes which constitute the building blocks of *dali/da li* are combined.

   In a (2016) paper on the different pragmatic types of focus, Bianchi, Bocci and Cruschina suggest that while the position in which mirative and corrective foci are interpreted is always the same, there is optionality concerning what copy one may spell out. More specifically, they postulate that mirative and corrective

---

23.   Jula is spoken in parts of Burkina Faso, Mali and Côte d'Ivoire.

foci must always move to the specifier of a left-peripheral Focus projection, as it is only in this position that the specific conventional implicature associated with these types of foci can be licensed. Nothing however prevents one from giving overt spell-out to the copy at the bottom of the movement chain, essentially undoing the effect of the movement derivation at PF. According to Bianchi, Bocci and Cruschina, this captures the fact that in structures like (48) the focus may either front or remain in situ, with no apparent consequences on the meaning of the sentence, the idea being that the focus is fronted in either case:

(48)   a.   I saw JOHN (not PAUL)!
       b.   JOHN I saw (not PAUL)!

In this paper, I would like to resort to a similar analysis to capture the optionality exhibited by group I languages concerning the possibility of having a focus either precede or follow PolC. In particular, I would like to suggest that there is optionality concerning which of the two copies of *dali/da li* may be given overt spell out: either the one at the top of the movement chain (=the head of the Int projection), or the one in the head of the Focus projection (=where *da* and *-li* are combined).

Now that we possess a derivation for *dali/da li* sentences containing a fronted focus, we are also in a position to account for why the β order is only an option in the specific subgroup of languages we have identified. According to the analysis we have been developing, the reason why group I languages display the possibility of having FOC locally precede PolC is because, at some point in the derivation of a focused polarity question, PolC is in a position linearly following the specifier of FocP, where the fronted focus is to ultimately surface. This is due to the morphologically complex nature of PolC in these languages: in group I languages, PolC results from the combination of an enclitic [+ focus] morpheme together with a proclitic morpheme specified as [−realis], a process which provides a phonological host for both clitics.

On the other hand, the polarity complementizers found in group II languages are not equally morphologically complex. These are *če* (Slovene), *chy* (Ukrainian), *czy* (Polish), *či* (Slovak). Also consider *se* (Italian) and *of* (Dutch), which, although not strictly part of group II languages, still pattern with these in exhibiting only the α order. An argument could be made for Czech *jestli*, which is clearly morphologically heavier than its counterparts in group II. As we already saw in section III., however, *jestli* originates from *jest*, a verbal form no longer in use. Crucial in our analysis of MAC, BUL and BSC *dali/da li* is the fact that the marker *da* is still used in present-day language as an independent, fully functional subordinating conjunction/subjunctive marker; this legitimizes an analysis of *dali/ da li* as resulting from the union of two independent building blocks. In the case of Czech *jestli*, on the other hand, we likely have a fully grammaticalized type of

polarity complementizer which is taken from the lexicon directly in the form in which it ultimately surfaces.

For Slovene *če*, Ukrainian *chy*, Polish *czy*, Slovak *či* as well as for Czech *jestli*, I thus adopt the same analysis suggested by Rizzi (2001) for Italian *se:* in these languages, the interrogative complementizer is generated directly in the head of IntP. Since PolC is generated directly in Int, at no point in the derivation of an embedded polarity question with a fronted focus will PolC be found in a position *preceding* the focalized constituent; this is because PolC will always be externally merged only after the focus has completed all of the movement steps associated with the focus fronting process. As such, we expect that the only way for a focus to linearly precede PolC in group II languages – and thus for the β order to obtain– is by fronting the focus to a CP higher than the one in which the base-generated PolC is merged. Indeed, this is precisely what we observe in group II languages: β is only possible as part of a non-local configuration.

Recall that the original observation that BUL can exhibit both the β and the α order is Ilyana Krapova's (Krapova 2002). The analysis I have just presented in similar in spirit to the analysis that Krapova herself suggests, but also crucially different. Krapova (2002) suggests that PolC is generated in a position lower than Focus, and may optionally raise to a position above it. She then postulates that optional movement of PolC accounts for languages like Bulgarian, which displays optionality between α and β, whereas mandatory movement of PolC to this pre-Focus position accounts for languages like Italian, where only α is observed. My analysis patterns with Krapova's in accounting for the existence of the β order by assuming movement of PolC, but unlike hers, it takes PolC to be moved only in a specific subset of languages (those that display the optionality): in all others, PolC is base-generated directly above Focus. My analysis then links a movement derivation of PolC to the specific morphological make-up of PolC itself: in group II languages, movement arises because PolC results from two independent clausal markers that need to be combined. Unlike Krapova's, my analysis provides a specific frame to understand why optionality arises in only some Slavic languages. Moreover, by stipulating that β is the exception rather than the norm, it accounts for the reduced availability of β in the linguistic sample examined in this paper.

5.3   The reasons behind alternative spell-out

The derivation in (42), coupled with Bianchi et al.'s (2016) alternative spell-out hypothesis, gives *dali/da li* means and opportunity to surface lower than the fronted focus. Neither our purported derivation nor the alternative spell-out hypothesis does however explain why Macedonian, Bulgarian and BSC speakers would ever choose to spell out an intermediate copy of *dali* rather than the top-

most one. We also have yet to account for why the order FOC < PolC seems to be less acceptable in BSC than in MAC and BUL.

While developing the generalization in (33) – which allowed us to single out group I languages–, we noticed that those languages which allow for FOC<PolC also have the option of realizing POL as -li. When POL is realized as -li, POL must follow the constituent in focus by virtue of the enclitic nature of -li; this automatically generates the order FOC<POL. On the assumption that both *dali* and -li are instances of POL, the existence of the -li option offers us with a straightforward explanation of why spelling out the intermediate copy of *dali* is possible in group I languages: it is because the order FOC<POL is already a grammatical ordering configuration in these languages.

The very same logic provides us with a potential explanation for why β is less common in BSC than it is in MAC/BUL. Consider Table 2, which I take from Schwabe (2004:10);[24] this table details acceptable enclisis sites for-li in those Slavic languages which feature this morpheme. In Table 2, a full circle indicates productive usage, whereas an empty circle indicates archaic usage.

As Table 2 shows, Bulgarian and Macedonian speakers productively use -li with both verbal heads and XPs, in both matrix and embedded polarity questions. Speakers of BSC also have the possibility of having -li encliticize onto a DP/ PP, but unlike in Macedonian and Bulgarian, this usage is perceived as archaic, by some speakers at least. According to Rudin et al. (1999) (see also references therein), XP-*li* in modern BSC is not accepted by all speakers and is limited to single-words XPs.

Therefore, while all speakers of MAC and BUL speakers accept XP-*li*, only some speakers of BSC do. Differently put, while for some BSC speakers FOC<POL is an acceptable configuration, for some other BSC speakers it is not. Consequently, whereas all MAC/BUL speakers always have a compelling reason to spell out the intermediate copy of PolC (i.e., the resulting configuration, FOC<POL, is already part of their grammars), not all BSC speakers do.

6.   *Dali da, *dali li* and other sequences of clausal markers

In this section, we focus on what possible combinations of different clausal markers are grammatical in the three languages which are the object of our analysis; we

---

24. Note that Schwabe's original table is more detailed than the one I have provided in Table 2, as hers features an additional Slavic language (Sorbian, spoken in Saxony and Brandenburg – Germany –, by around 50.000 speakers). Schwabe's table also details the usage of -li in conditional clauses, which I have omitted in that not strictly relevant for the purposes of this paper.

Table 2. Grammatical enclisis sites for-*li* in Slavic languages

		V-li		XP-li	
		Matrix	**Subord.**	**Matrix**	**Subord.**
SS	BSC	•	•	○	○
	BUL	•	•	•	•
	MAC	•	•	•	•
	SLO				
WS	CZ		•		○**
	SLK				
	POL	○			
ES	RUS	(•)	•	(•)	•
	UKR				
	BEL				

* Note that Czech is also marked as having the possibility of encliticizing -*li* onto XPs, although only in subordinate clauses and in a non-productive way. None of my Czech informants however accepts this usage of –li; this is also supported by Toman (1996), who reports that -*li* structures are only limited to V-*li*. This is perhaps due to regional variation. I have thus marked the corresponding cell with an asterisk.

will see that one such particular combination, *dali da*, warrants a reconsideration of our analysis of *da*.

A distinctive trait of Slavic languages is the possibility of "piling up" different clausal markers, as can be seen from the Slovene example in (49). On top of a *wh*-word, (49) features also the polarity complementizer *če* and the irrealis marker *da*:

(49)   *Kdo če da pride?*   (SLV)
      Who if da comes?
      'Who is said to be coming?'

                                  (Hladnik 2010:15)

Likewise, both -*li* and *da* may be found in combination with other elements already marking the clause as being interrogative or *irrealis*, in all the languages under discussion. -*li* may for instance combine with a *wh*-word to give rise to what Rudin (1986) refers to as "emphatic" questions, and which Hauge (p.c.) describes

as questions which are not truly information-seeking, either because the answer is obvious/understood, or because no real answer exists:[25, 26]

(50) Но кой ли има време за това?   (BUL)
      No koy li ima vreme za tova?
      But who -li has time for that?
      'But WHO has time for that?'[27]

<div align="right">(From The Bulgarian National Corpus)[28]</div>

Obviously, not all possible combinations of different clausal markers are acceptable; thus, we want to make sure that our model can correctly rule out those combinations which are impossible.

In its current formulation, the derivation in (42) correctly predicts that sequences of the type of '*dali* (FOC) *li*' should be impossible:[29]

(51) * Чудя   се дали книгите   ли Иван ще купи   (BUL)
      Chudja se dali KNIGITE   li Ivan shte kupi
      I-wonder refl. if   BOOKS.THE li Ivan will buy

The ungrammaticality of (51) follows if we assume that the *li* found in the morphologically complex word *dali/da li* is indeed the same *li* we find in questions where this is the only element signaling that the clause is a yes/no interrogative:

---

25. See Kaspar (2017) on how *že* performs a similar function in Czech.

26. The specific pragmatic interpretation associated with these types of questions follows from our analysis of *-li* as focus marker (see section V.I.I). If *-li* forces a focus interpretation of the constituent which is fronted to a pre *-li* position, it follows that the interpretation of (52) is roughly equivalent to English "but WHO has time for that(?)". Here, emphatic stress on the interrogative constituent could be used to convey that the existential implicature generally associated with *wh*-elements must be canceled, as presumably no entity x exists such that x has enough time.

27. Thanks to Kjetil Rå Hauge for suggesting this example to me.

28. Freely accessible at http://dcl.bas.bg/bulnc/en.

29. The only case in which *-li* is allowed to co-occur with *dali* are echo questions of the type of (i) below, i.e. when one is requesting confirmation for a question recently uttered:

(i)   A: 'Is she at home?'
      B: Дали е   в къщата ли? Не знам.
          Dali e   v kŭshtata li? Ne znam
          Dali AUX in house?   Not I-know
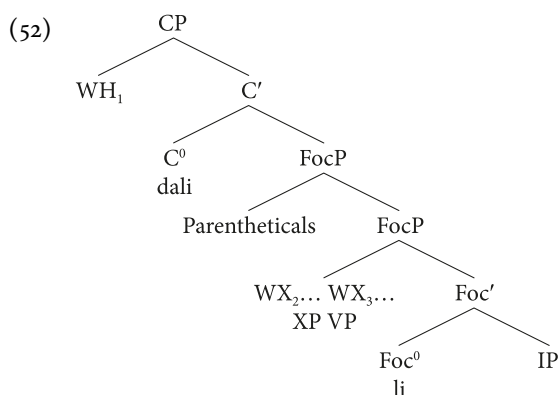   'If she is at home? I don't know'

<div align="right">(Rudin 1985: 65)</div>

These are special cases in that the clause containing *dali,* i.e. the whole sentence preceding *-li* is treated as a single unit, almost as if it were a simple DP (see Rudin 1985).

in both cases, this is generated in the head of FocP. This automatically accounts for the fact that *dali/da li* the complementizer is incompatible with *-li* the enclitic morpheme.

In this respect, (42) has an advantage on Dukova-Zheleva's (2010) purported structure of the left periphery, as hers cannot automatically derive the incompatibility of *dali/da li* with *-li* the way (42) does. Dukova-Zheleva suggests that, while *-li* is merged as the head of FocP, *dali* is merged in the specifier of a higher C projection, as illustrated in (52):

(52)

```
                    CP
                  /    \
               WH₁       C′
                       /    \
                     C⁰      FocP
                     dali   /    \
                  Parentheticals  FocP
                              /        \
                        WX₂… WX₃…       Foc′
                          XP VP       /    \
                                   Foc⁰     IP
                                    li
```

(Dukova-Zheleva 2010: 55)

If (52) were correct, however, we would expect sentences like (51) to be acceptable, a scenario that Dukova-Zheleva is forced to rule out independently by suggesting that the *wh-* feature on *-li* is incompatible with the *wh-* feature on *dali*, since the clause only needs to be typed as being [+interrogative] once. Moreover, we saw that *dali* and *-li* perform partially overlapping functions in that the presence of either signals that the relevant clause is to be interpreted as a polarity question. This naturally follows from (42) since what we are suggesting is that the derivation of *dali* and *-li* questions is at least partially overlapping: a focus marker (*-li*) is involved in both cases and is crucially generated in the same position in both structures. On the other hand, the functional equivalency of *dali* and *-li* does not immediately follow from (52).

If the derivation in (42) has no problem accounting for the ungrammaticality of sequences of the type of '*dali li', it faces some apparent issues in dealing with the fact that sequences of the type of '*dali da*' are not simply attested, but are also very frequent. Two examples of the relevant structure are provided below. In (53) we have a matrix yes/no interrogative, in (54) an embedded one:

(53) Da li da Vesna pročita ovu knjigu?   (BSC)
      Da li da Vesna reads   this book
      'Should Vesna read this book?'[30]

                              (Todorović & Wurmbrand 2017: 7, from Vrzić 1996: 292)

(54) Теон неволно   се   зачуди   дали да изрече   молитва   (BUL)
      Teon nevolno   se   zachudi   dali da izreche   molitva
      Teon unintentionally REFL. wondered dali   da pronounce prayers
      'Theon unintentionally wondered whether to recite prayers'

                                               From *The Bulgarian National Corpus*

In both (53) and (54), the addition of a second *da* morpheme results in the polarity question acquiring a deontic undertone. This is particularly evident in (53), which the authors decided to translate to English by resorting to the modal 'should', but is none the less present in embedded structures like that in (54): even in English, a possible paraphrase of 'Theon wondered whether to recite prayers' could be, depending on the context, 'Theon wondered whether he *should* recite prayers'.

It is not so much the existence of sequences of the type of '*dali da*' which is problematic, but rather the fact that there is a clear difference in meaning between plain *dali* polarity questions, and *dali da* ones: the deontic undertone which characterizes the latter structures is not observed in the former. Accordingly, it would be difficult to claim, for instance, that the extra *da* in *dali da* is nothing but the spell-out of the external merge site of *da*, i.e., the foot of *da*'s movement chain. If that were the case, we would expect the meaning of *dali da* structures to be identical to that of plain *dali* ones, or at least for the meaning not to change so obviously.

The key to solving the puzzle, I believe, lies in fully acknowledging the extreme flexibility displayed by the particle *da*. Recall from section V that *da* is quite the wild card of the Balkan language family, its functions ranging from the expression of deontic modality, to its key role in licensing optatives, nonfactual subordinate clauses and counterfactuals. Accordingly, I will be arguing that whereas *da* always spells out [−realis], the exact compositional import of this [−realis] feature hinges on where *da* is merged exactly, resulting in the possibility of having multiple *das* being merged in different points of the structure. Underlying the string *dali da* are then two distinct irrealis markers, each performing a different function.

---

30. Translation as in Todorović & Wurmbrand (2017). Other authors (e.g., Rudin (1990) and Dukova-Zheleva (2010) for Bulgarian) translate even matrix *dali da* structures simply by using 'whether to'.

Let us start by considering Todorović & Wurmbrand's (2017) account of the role and distribution of *da* in BSC. Todorović & Wurmbrand treat *da* as a finiteness "visualizer", suggesting that its function is that of overtly spelling out the [+finite] feature of any finite verb onto which *da* procliticizes. This description of *da* is perhaps a bit misleading in the context of this paper, where I have suggested that *da* is employed in MAC and BUL to realize what would be expressed with an infinitival structure in languages like English.[31] It however makes more sense when describing BSC in that, unlike BUL and MAC, at least some varieties[32] of BSC have retained a system of grammaticalized infinitives. These represent a second strategy to form infinitival-like structures in BSC, the other one being the same '*da* + present tense' structure we also observe in BUL and MAC. The two strategies are illustrated in (55):

(55) a. *Odlučila sam   da prevodim   pesmu*   (BSC)
          Decided aux.1.sg da translate.1sg.pres poem
      b. *Odlučila sam   prevedem   pesmu*
          Decided aux.1.sg translate.inf poem
          'I decided to translate the poem'

(Todorović & Wurmbrand 2017: 2)

(55a) has the same meaning of (55b), but crucially the embedded verb is here formally [+ finite]. It is in this sense that *da* is a finiteness visualizer: it is followed by verbs which are always grammatically [+ finite].

Todorović & Wurmbrand suggests that *da* overtly spells out the finiteness of a verb whenever no other feature or clausal marker is present to perform that same function. Crucially, this means that *da* may be realized more than once within the same sentence, depending on the number of verbs that the sentence features. The authors also suggest that *da* is capable of spelling out different functional heads (C, T or little v), depending on the verb whose finiteness *da* goes to spell out. Following Wurmbrand (2001, 2014, 2015), and on the basis of phenomena like clitic climbing,[33] Todorović & Wurmbrand argue in particular that different predicates select for complement clauses of different sizes: CPs, TPs or vPs. For instance, whereas predicates like *decide* select for TPs, verbs like *try* select for vPs, as suggested by the fact that the clausal complement of a verb like *decide* may contain time references, whereas that of *try* may not:

(56) a.   She decided to eat tomorrow

---

31. See again section V.

32. See again footnote 14.

33. On the assumption that CPs are barriers for such a phenomenon, see Wurmbrand (2015).

b.   She tried to eat *(tomorrow)

In BSC, both (56a)–(b) could be expressed with a *da* structure, but crucially the *da* morpheme would then be spelling out two different clausal layers: the TP in (56a), and the vP in the case of (56b).

Example (57) illustrates a structure where more than one *da* appears within the same clause:

(57)   *Kaže da će   da dođe.*
        Says da will.3.sg da come.3.sg.pres
        'He says he will come.'

(Todorović & Wurmbrand 2017, from Sočanac 2011: 55)

(57) features the auxiliary *će* and the finite verb *dođe*. Todorović & Wurmbrand suggest that whereas the higher *da* spells out C, the second, lower *da* spells out v, as illustrated in (58):

(58)
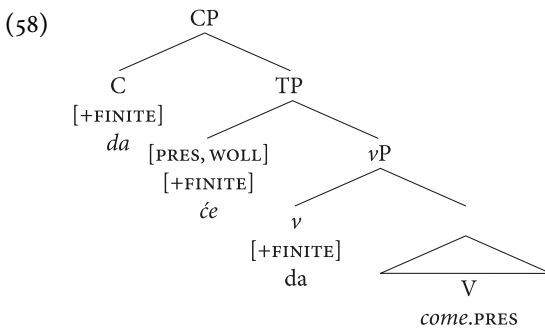
(Todorović & Wurmbrand 2017: 10)

The description of *da* as a finiteness visualizer may work for languages like BSC, which have retained a grammaticalized system of infinitives, but falls somewhat short when accounting for languages like BUL and MAC, which have not. In this paper, I will be following Todorović & Wurmbrand (2017) in assuming that *da* may occur more than once within the same clause, and that when it does, the different *da*s goes to spell out different functional projections. Unlike Todorović & Wurmbrand, however, I will assume as I have already done above that *da* overtly spells out the presence of a [−realis] feature, and not that of a [+finite] one.

Now that we are familiar with the mechanics of *da* and multiple *da* structure, let us go back to the question of how to account for strings of the type of *dali da*. If, as argued in Todorović & Wurmbrand (2017), it is possible for *da* to spell out different clausal layers, a way of accounting for strings of the type of *dali da* emerges: whereas the *da* we see as spelled out in the morphologically complex word *dali/da li* originates within the CP, the lower *da* spells out the edge of a lower clausal layer,

namely the TP or (if we are to follow Todorović & Wurmbrand) the vP depending on the specific predicate. The lower *da* would procliticize onto the associated verb, whereas the higher *da* would be forced to procliticize onto *-li*, the closest available phonological host which is also the only one available in C given the lack of a verbal host in this layer.

The suggested analysis saves us from our current predicament in that it accounts for the presence of a second, lower *da* in strings of the type of *dali da*, but at this point is still highly speculative. In particular, if T-layer *da* already spells out [−realis] for that specific clause, what reason would there be to also do so at the level of the CP?

The crucial idea I will be adopting is that merging a [−realis] morpheme at the level of the CP results in an effect which is substantially different from doing so in a lower projection, which justifies merging two distinct *da* morphemes within the same clause.

How does C-layer *da* differ from T-layer *da*? To answer that question, let us consider once more the difference between embedded *dali* questions on the one hand, and embedded *-li* questions on the other. In section IV. we saw that whereas *dali* is the most pragmatically neutral option to form an embedded interrogative, embedded *-li* questions are perceived to convey something akin to direct quotation or free indirect speech. If C-layer *da* does indeed perform a function which is different from lower *da*, then logic dictates that, whatever this function is, it should result in preventing the embedded clause from having its own illocutionary force. But why would [−*realis*] be connected to the presence or lack of independent illocutionary force?

According to Palmer (2001), the indicative/subjunctive distinction on the one hand, and the use of various realis and irrealis particles on the other, should both be treated as instantiations of the same phenomenon, namely the expression of the concepts of *Realis* and *Irrealis*.[34] Cross-linguistically, we see that the subjunctive mood is often used to mark subordination, and to prevent a subjunctive predicate from being interpreted within the scope of illocutionary force of the matrix clause. The latter case is exemplified by Belhare (Tibeto-Burman), as illustrated in (59). In this example, the subordinate clause *rakharana*[35] is not in the subjunctive mood. According to Bickel (1993), since it is not subjunctive, the interpretation of this clause is ambiguous, the possible interpretive differences relating to the degree of integration of said dependent predicate with the matrix:

---

34. In Palmer (2001), "Realis" capitalized refers to the concept, whereas "realis" non-capitalized refers to the specific morphological realization of such concept, for instance through the use of realis particles.

35. Belhare has polysynthetic morphology.

(59)   *Rak-khar-a-na   hab-he   i?*   (BYW)
       Get_tired-TEL-CONJ-TOP weep-PT Q?
       a.   'Did he cry because he was tired?' (illocution attraction)
       b.   'When he was tired, did he cry?' (no illocution attraction)

                                                         (Bickel 1993: 33)

Bickel describes (59a) as an instance of "illocution attraction": the dependent predicate remains in the scope of the interrogative illocutionary force associated with the matrix predicate. In (59b), on the other hand, the dependent predicate retains independence from the illocutionary force of the main predicate.

If subjunctive mood and [−realis] particles are indeed instantiations of the same phenomenon, we can devise an explanation for the differences between *dali* and *-li* questions which operates on a logic similar to that suggested for (59). Specifically, we may suggest that the [−realis] marker *da* merged in CP acts as a barrier for illocutionary force in ways similar to what subjunctive mood does in Belhare. The idea is the following: in an embedded *dali* question, *da* acts as a blocker, preventing the interrogative feature on the embedded polarity question from reaching the matrix left periphery and thus from typing the entire sentence as being interrogative. In embedded *-li* questions, on the other hand, since no *da* is present, no such blocker of illocutionary force is active; the interrogative feature on the embedded clause becomes fully activated interrogative illocutionary force. Since a different illocutionary force is already present at the level of the matrix CP, preventing the interrogative force of the embedded from percolating all the way up to the matrix, the sentence then ends up being interpreted as featuring two different illocutionary forces, each applying at different clausal levels.

Recall that *dali/da li* may also introduce matrix polarity questions. In matrix polarity questions, since there is no clause higher than the one in which *dali/da li* is merged, the presence of *da* in the C layer in inconsequential: there is no need to block the interrogative feature on that clause from typing as interrogative any higher clause. This generates optionality with respect to whether the matrix polarity interrogative may be formed through the use of *dali*, or of *-li* alone.

What about *da* merged in T, or lower? T-layer *da* works by removing the indicative component from (i.e. by imposing a subjunctive reading on) the event being described by the predicate. This causes such an event to no longer be epistemically evaluated against the actual world, as illustrated by the contrast between (60a) and (60b). Ultimately, this is what is responsible for the emergence of the deontic undertone we have observed in *dali da* structures, as one can gauge by the English translations of the examples below.

(60) a. Попита   ме дали е   отворила прозореца.
          Popita   me dali e   otvorila   prozoretsa
          (S)he-asked me dali AUX opened   the.window
          'S/he asked me whether s/he has opened the window'

      b. Попита   ме дали да отвори прозореца.
          Popita   me dali da otvori   prozoretsa
          (S)he-asked me dali   da opened the.window
          'S/he asked me whether to open the window'

T-layer da always works by removing the indicative component from an event, but the way this is actualized also depends on whether or not *da* is selected by a higher verb. A particularly compelling example in this respect are structures of the type of '*da* Verb *li* (Verb)'; these also illustrate how the interpretation of *da* is entirely dependent on its extraction site, not on its surface position. Recall that *-li* is morphologically enclitic and thus triggers the fronting of a finite verb (or of a constituent in narrow focus if present) to its specifier: this is so that *-li* may have a phonological host onto which to encliticize. If only one verb is present, and *da* is present, the polarity question acquires the same deontic undertone we observed in (61) in conjunction with the presence of *dali*:

(61) Да почукам ли?   (BUL)
      Da pochukam li?
      Da I-knock   -li?
      'Should I knock?'

We can assume the following derivation for (61): the verb *pochuka* is externally merged in *v,* and *da* is merged as the head of T. The left periphery is then created and *-li* is externally merged as the head of FocP; this operation also triggers the fronting of the entire *da* + verb complex to its specifier:

(62) [$_{CP}$ [$_{FOC}$ da pochuka [$_{FOC°}$ -li [$_{TP}$ [$_{T}$ da [$_{vP}$ [$_{v}$ pochuka]]]]]]]

Consider now the following example, where a second verbal structure follows *-li.* As the reader can gauge from the translation, the deontic undertone has now been replaced by a pure polarity-question interpretation:

(63) Да злорадстваш ли дойде?   (BUL)
      Da zloradstvash   li doyde?
      Da you-gloat   li you-came
      'Have you come to gloat?'

                                  Example taken from *The Bulgarian National Corpus*

(61) and (63) are at least partially string-identical (up to '*da V li*') but their derivation is crucially different. More specifically, what is fronted to a pre- *-li* position in (63) is not the closest available verb, but its subordinate predicate:[36]

(64) $[_{CP1}$ $[_{FOC}$ da zloradstvaš $[_{FOC°}$ -li $[_{TP}$ doĭde $[_{CP2}$ $[_{TP}$ $[_{\overline{T}}$ ~~da~~ $[_{vP}$ $[_{v}$ ~~zloradstvash~~]]]]]]]

What is fronted in (63) is the *irrealis* particle which goes to remove the indicative component from a subordinate verb. Even though an irrealis marker surfaces sentence-initially in both (61) and (63), its import on the overall structure is then crucially different.

The analysis I have suggested for *da* accounts for a number of equivalences we have seen above. For example, we saw that matrix *dali da* structures (e.g (53)) are equivalent in meaning to matrix *da V li* ones (see (61)): both can be translated as deontic polarity questions. At first sight, this is surprising: according to the analysis we have developed above, *da V li* structures effectively feature one irrealis morpheme less than *dali da* ones. Compare also with embedded domains, where *dali da* questions do differ in meaning from plain *dali* ones.

The *da V li*/matrix *dali da* equivalence holds because the extra *da* which goes to combine with *li* to form *dali* in *dali da* structure is optional. The derivation for *da V li* structures was already presented in (62) for the clause *da pochuka li?* ('Should I knock?'): I repeat it in (65). Compare it with the derivation of the equivalent *dali da* structure, *dali da pochuka?* (also 'Should I knock?'), which I present in bracket notation in (66):

(65) $[_{CP}$ $[_{FOC}$ da pochuka $[_{FOC°}$ -li $[_{TP}$ $[_{\overline{T}}$ ~~da~~ $[_{vP}$ $[_{v}$ ~~pochuka~~]]]]]]]

(66) $[_{CP}$ $[_{FOC°}$ da-li $[_{Fin}$ ~~da~~ $[_{TP}$ $[_{T}$ da $[_{vP}$ $[_{v}$ pochuka]]]]]]]

---

36. It is of course also possible to form a matrix polarity question by fronting the higher verb instead. The following two examples (both taken from the *Bulgarian National* Corpus), for instance, are equally acceptable:

(i)   искаш   ли да умра?
      Iskash   li   da umra?
      You-want li   da I-die
      'Do you want me to die?'

(ii) Да умра ли искаш?
      Da umra li   iskash?
      Da I-die li   you-want?
      'Do you want me to die?'

Recall that we have hypothesized that C-layer *da* goes to stop the illocutionary force of the clause it is merged in from migrating to higher CPs. Since (65) is already a matrix CP, whether C-layer *da* is present or not is inconsequential for the transmission of illocutionary force: there is no higher clause to which such force may be transmitted. What C-layer *da* can do in these cases is provide a phonological host for *li*, so that the verb need not front to do so itself. If the verb does front, as in (66), it will drag along T-layer *da*, as this *da* needs a host as well and thus cannot remain in T. Since this *da* was extracted from T and not from C, we are effectively fronting the irrealis morpheme involved in the removal of the indicative component from the verb *pochuka*. Even if this *da* ultimately surfaces sentence-initially, its function is not that of C-layer *da*. The equivalence between (65) and (66) thus follows.

7.   Conclusions

The goal of this article was to investigate the relative distribution of two types of left-peripheral elements: fronted constituents in narrow focus, and polarity complementizers (PolC). I have focused on accounting for restrictions observed at the *local* level, namely whenever these two elements surface in an identical left periphery.

At the local level, several languages only exhibit one possible ordering configuration: the polarity complementizer can only *precede* a fronted constituent in narrow focus (configuration which I have labeled *alpha*). In these languages, the *beta* order (i.e., focus preceding PolC) is locally never possible.

Contra Abels (2012), I have argued that the reason why the β configuration is locally ungrammatical in some languages cannot be due to restrictions on the extraction of a focus across an interrogative complementizer. If that were the case, we would expect that the long-distance movement of a focus across PolC would be equally ungrammatical, whereas this is never the case.

Not all languages disallow foci from locally preceding PolC. In the Slavic subgroup, for example, two (three if we count Bosnian-Serbian-Croatian, for which acceptance of β seems less widespread) out of the eight languages which possess a free morpheme to realize POL allow for β locally: these are Macedonian and Bulgarian. This raises the question of why an ordering configuration which appears to be cross-linguistically disfavored is possible in precisely this group of languages.

The fact that β may be possible locally in these specific languages, I have argued, is no accident, but rather a consequence of the morphologically complex nature of PolC in the three languages under discussion. In MAC, BUL and BSC, PolC is spelled out as *dali/da li*. Crucially, both of the two morphological building

blocks which form this word, *-li* and *da,* are fully functional clausal markers in present-day MAC, BUL and BSC, and may be used independently of each other. I have argued that a movement-plus-incorporation process underlies the derivation of PolC in these languages; *dali/da li* is derived through the incorporation of two distinct functional morphemes. One is *da*, a multi-functional [−realis] marker which is proclitic in nature. The second element is the structurally higher [+ focus] marker *-li*, an enclitic morpheme. This incorporation process provides a phonological host for both clitics. In all the languages where the α order is the only option locally, on the other hand, the morpheme which lexicalizes POL is either morphologically simple or no longer transparent, and thus does not result from the incorporation of distinct clausal markers.

Crucially, because a movement-plus-incorporation analysis underlies the derivation of PolC in Macedonian, Bulgarian and BSC, a copy of PolC itself is present in a position lower than where the focus is internally merged when fronted. The presence of this lower copy generates optionality with respect to what copy of PolC is given overt spell-out: when one spells out the copy of *dali/da li* found where the incorporation of *da* with *-li* takes place, the β order ensues. When the head of the movement chain of PolC is pronounced instead, the α order is obtained. We have also seen that motivation for spelling out an intermediate copy of *dali* comes from the fact that β- group languages already possess a structure which results in the order FOC < POL: XP *-li* structures.

In the last sections of this article, I have argued that the proclitic morpheme *da* always spells out *irrealis*, but the specific compositional import of this *irrealis* component is dependent on where *da* is merged exactly. Factors which affect the nature of *da*'s irrealis import are whether the verb is selected by a higher predicate, and in which specific functional projection (C, or anything lower) *da* is merged exactly.

Acknowledgements

References

Abels, Klaus. 2012. The Italian left periphery: A view from locality. *Linguistic Inquiry 43*(1). 229–254. https://doi.org/10.1162/LING_a_00084

Bianchi, Valentina, Bocci, Giuliano & Silvio, Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics 9*(3). https://doi.org/10.3765/sp.9.3

Bickel, Balthazar. 1993. Belhare subordination and the theory of topic. *Studies in clause linkage*, 23–55.

Callegari, Elena. 2018. Understanding Word Order in the Left Periphery. University of Oslo. (Doctoral dissertation).

Callegari, Elena. 2021. *Focus Architecture in Manding Language*. Handout, Incontro di Grammatica Generativa (IGG), Siena, 2021.

Cardinaletti, Anna. 2008. On different types of clitic clusters. *The Bantu-Romance Connection*, 41–82. https://doi.org/10.1075/la.131.06car

Chomsky, Noam. 1970. Deep structure, surface structure, and semantic interpretation. In R. Jakobson & S. Kawamoto (eds.), *Studies in General and Oriental Linguistics*. Tokyo T.E.C. Corporation.

Cruschina, Silvio. 2010. Syntactic extraposition and clitic resumption in Italian. *Lingua 120*(1). 50–73. https://doi.org/10.1016/j.lingua.2009.04.002

Cruschina, Silvio. 2016. *Licensing Focus Movement: Distributional and Interpretive Properties*. Handout, 1st SynCart workshop, 11–15 July 2016.

Dukova-Zheleva, Galina. 2010. Questions and focus in Bulgarian. University of Ottawa. (Doctoral Dissertation).

Englund Dimitrova, B. 1977. Yes/no-questions in Bulgarian and Macedonian: form. Almqvist & Wiksell international. (Doctoral dissertation).

Friedman, Victor. A. 1993. Macedonian. In B. Comrie & G.C. Corbett (eds.), *The Slavonic Languages*. Routledge, London, pp. 249–305.

Friedman, Victor. A. 2018. Parallel Universes and Universal Parallels: Balkan Romani Evidential Strategies. *Balkan Syntax and (Universal) Principles of Grammar 285*(37) https://doi.org/10.1515/9783110375930-003

Halliday, M.A.K. 1967. *Intonation and grammar in British English*. The Hague: Mouton. https://doi.org/10.1515/9783111357447

Hansen, Björn, Letuchiy, Alexander & Izabela Błaszczyk. 2016. Complementizers in Slavonic (Russian, Polish, and Bulgarian). *Complementizer Semantics in European Languages 57* (175). https://doi.org/10.1515/9783110416619-008

Hauge, Kjetil. R. 1999. The word order of predicate clitics in Bulgarian. *Journal of Slavic linguistics 7*(1). 89–137.

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language 10*. 41–53.

Hladnik, Marko. 2010. Restrictive relative clauses in Slovene and its spoken varieties. Utrecht: Utrecht University. (Master's thesis).

Hock, Hans H. 1992. What's a nice word like you doing in a place like this? Syntax vs. phonological form. *Studies in the Linguistic Sciences 22*(1). 39–87.

Jackendoff, Ray. S. 1972. *Semantic interpretation in generative grammar*. Cambridge, MA. MIT Press.

Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general and historical linguistics* (Vol. *1*). Cambridge University Press.

Kaspar, Jiri. 2017. Syncretism: A case study of the particle že in Czech. UCL (University College London). (Doctoral dissertation).

King, Tracy H. 1994. Focus in Russian yes-no questions. *Journal of Slavic linguistics*, 92–120.

Krapova, Iliyana. 2002. On the Left Periphery of the Bulgarian sentence. *University of Venice Working Papers in Linguistics 12*. 107–128.

Kochovska, Slavica. 2010. *Macedonian direct objects, clitics and the left periphery*. Rutgers The State University of New Jersey-New Brunswick.

Legendre, Géraldine. 2000. Morphological and prosodic alignment of Bulgarian clitics. *Optimality theory: Syntax, phonology, and acquisition*, 423–462.

Izvorski, Roumyana. 1993. On wh-movement and focus movement in Bulgarian. *Second Conference of the Student Organization of Linguistics in Europe (CONSOLE), University of Tübingen.*

Lindstedt, Jouko. 2010. Mood in Bulgarian and Macedonian. *Mood in the languages of Europe*, 409–421. https://doi.org/10.1075/slcs.120.23lin

Mihaliček, Vedrana. 2012. Serbo-Croatian word order: A logical approach. The Ohio State University. (Doctoral dissertation).

Nikov, M. D. 1976. Leksiko-sintaktichna struktura na istinskite obshti vaprosi vav frenski i balgarski ezik, *Byuletin za sapostavitelno izsledvane na balgarski ezik s drugi ezitsi 3*, 115–34.

Palmer, Frank. R. 2001. *Mood and modality.* Cambridge University Press. https://doi.org/10.1017/CBO9781139167178

Radanović-Kocioć, Vesna. 1988. The grammar of Serbo-Croatian clitics: A synchronic and diachronic perspective. Urbana-Champaign: University of Illinois. (Doctoral dissertation).

Richards, Norvin. 2010. *Uttering trees.* Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/9780262013765.001.0001

Richards, Norvin. 2016. *Contiguity theory.* Cambridge, MA: MIT Press.

Rivero, Marai L. 2005. Topics in Bulgarian morphology and syntax: a minimalist perspective. *Lingua 115*(8). 1083–1128. https://doi.org/10.1016/j.lingua.2004.02.006

Rizzi, Luigi & Giuliano Bocci. 2017. Left Periphery of the Clause: Primarily Illustrated for Italian. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1–30. https://doi.org/10.1002/9781118358733.wbsyncom104

Rizzi, Luigi. 1997. The fine structure of the left periphery. *Elements of grammar: Handbook in generative syntax*, ed. by Liliane Haegeman. 281–337. https://doi.org/10.1007/978-94-011-5420-8_7

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. *Current Studies in Italian syntax: Essays offered to Lorenzo Renzi*, 287–296.

Rizzi, Luigi. 2004a. Locality and left periphery. *Structures and beyond. The cartography of syntactic structures 3.* 223–251.

Rizzi, Luigi. (Ed.). 2004b. *The structure of CP and IP: The cartography of syntactic structures (Vol. 2).* Oxford University Press.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, 91–136.

Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics 1*(1). 75–116. https://doi.org/10.1007/BF02342617

Rooth, Mats. 1999. Association with focus or association with presupposition? In P. Bosch & R. van der Sandt (eds.), *Focus: Linguistic, cognitive, and computational perspectives*, 232–246.

Rudin, Catherine. 1983. DA and the category AUX in Bulgarian. *Linguistic categories: Auxiliaries and related puzzles*, 3–20. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-6989-6_1

Rudin, Catherine. 1986. Aspects of Bulgarian syntax. *Columbus: Slavica.*

Rudin, Catherine, Kramer, Christine, Billings, Loren, & Matthew Baerman. 1999. Macedonian and Bulgarian li questions: Beyond syntax. *Natural Language & Linguistic Theory 17*(3). 541–586. https://doi.org/10.1023/A:1006223423256

Schwabe, Kristine. 2004. The particle li and the left periphery of Slavic yes/no interrogatives. In H. Lohnstein & S. Trissler (eds.), *The Syntax and Semantics of the Left Periphery*. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110912111.385

Selvelli, Giustina. 2015. Su alcuni aspetti ideologici dei sistemi di traslitterazione degli alfabeti cirillici nei Balcani. *Studi Slavistici XII*. 159–180.

Shlonsky, Ur. 1988. Complementizer-cliticization in Hebrew and the empty category principle. *Natural Language & Linguistic Theory 6*(2). 191–205. https://doi.org/10.1007/BF00134229

Sočanac, Tomislav. 2011. Subjunctive in Serbian/Croation. *Geneva Generative Papers 7*, 49–70.

Szendrői, Kriszta. 2002. Stress-focus correspondence in Italian. *Proceedings of Going Romance 2000*, 287–305. https://doi.org/10.1075/cilt.232.16sze

Szendrői, Kriszta. 2003. A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review 20*. 37–78. https://doi.org/10.1515/tlir.2003.002

Szendrői, Kriszta. 2017. The syntax of information structure and the PF interface. *Glossa: a journal of general linguistics 2*(1). https://doi.org/10.5334/gjgl.140

Tabakowska, Elzbieta. 1997. Conceptualization: Conditionals as an instance of figure/ground alignment. *On conditionals again 143*. https://doi.org/10.1075/cilt.143.15tab

Todorović, N. & Suzi Wurmbrand. 2017. Suzi. 2017. Finiteness across domains. In T. Radeva-Bork & P. Kosta (eds.), *Current developments in Slavic Linguistics – Twenty years after*, 47–66.

Toman, Jindřich. 1996. A note of clitics and prosody. In A. Hallpern & A. Zwicky (eds.), *Approaching Second Position Clitics and Related Phenomena*. Stanford: ASLI Publications. 505–510.

Tomić, Olga M. 1996. The Balkan Slavic clausal clitics. *Natural Language & Linguistic Theory 14*(4). 811–872. https://doi.org/10.1007/BF00133364

Vakareliyska, Cynthia. 1994. Na-drop in Bulgarian. *Journal of Slavic linguistics*, 121–150.

Vrzić, Zvjezdana. 1996. Categorial status of the Serbo-Croatian "modal" da. In J. Toman *Annual Workshop on Formal Approaches to Slavic Linguistics: The college Park Meeting 1994*, 291–312. Ann Arbor, MI: Michigan Slavic Publications.

Wurmbrand, Suzi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: Mouton.

Wurmbrand, Suzi. 2014. Restructuring across the world. *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014*. Olomouc, 275–94.

Wurmbrand, Suzi. 2015. Restructuring cross-linguistically. *Proceedings of the North Eastern Linguistics Society Annual Meeting 45*. Amherst: GLSA, 227–40.

Address for correspondence

Elena Callegari
University of Iceland
Niels Treschows Hus
Niels Henrik Abels vei 36
Oslo 0315
Norway
ecallegari@hi.is

elena.callegari@ilos.uio.no
https://orcid.org/0000-0002-4928-7589

Publication history