The Relative Order of Foci and Polarity Complementizers: 
A Slavic Perspective

Abstract

According to Rizzi & Bocci’s (2017) suggested hierarchy of the left periphery, fronted foci (FOC) may never precede polarity complementizers (POL); yet languages like Bulgarian and Macedonian, where POL may also follow FOC, seem to provide a counterargument to such generalization. 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 POL is dependent on the morphological properties of the complementizer itself: in languages where the order FOC < POL is acceptable, POL is a complex morpheme derived through the incorporation of a lower functional head with a higher one. The order FOC < POL is then derived by giving overt spell-out to the intermediate copy of the polarity complementizer rather than to the highest one.

Keywords: Left Periphery, Fronted Foci, Polarity Questions, Complementizers, Bulgarian, Macedonian, Word Order.

I. 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 (POL), 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’

1 Thanks to Kjetil Rå Hauge, Yovka Tisheva and Victor Friedman for the helpful reviews and comments.
Concerning the relative order of these two elements, we have two logical possibilities: either
the constituent in focus will follow the polarity complementizer, or it will precede it. In this
article, I will 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) Order α: if < FOC (“se A TROMSØ”)
       Order β: FOC < if (“A TROMSØ se”)

According to Rizzi (2001) (see also Rizzi 2004, and Rizzi & Bocci 2017), foci and polarity
complementizers are associated with dedicated left-peripheral projections. In particular, the
complementizer “if” is said to be externally merged in the head of Int(errogaative)P, a
projection sandwiched in between two other projections equally specialized for hosting
complementizers, namely Force and Fin. Unlike IntP, however, Force and Fin may
exclusively host declarative complementizers. Rizzi’s suggested derivation for fronted foci is
slightly more complex, in that 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 a 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 situated lower than IntP, as illustrated in (3). (3) represents
the most updated version of the cartographic hierarchy of the left periphery:

3) Force < TopP < IntP < TopP < FocP < TopP < ModP < TopP < QEmb < Fin < IP
   (Rizzi & Bocci 2017, ex. 29)

Evidence in favor of the idea that FocP must be lower than IntP comes from grammaticality
patterns such as the ones illustrated in (4) below. 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”:

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

   2 At least at the time of writing (March 2019).
Based on the hierarchy in (3), we would then expect that the only way for a fronted focus to precede a polarity complementizer would be for the focus to be moved to a left periphery higher than the one in which POL is merged. In this case, the focus would linearly precede the polarity complementizer simply by virtue of having being dislocated to a different—and crucially higher—C-layer. This is illustrated in (5), where I use the notation “CP2” to indicate a higher left periphery:

5) \([\text{CP2} \ldots \text{FOC} \ldots [\text{CP1} \ldots \text{POL} < \text{FOC} \ldots [\text{TP} \ldots \text{FOC} \ldots]]]\)

In this paper, I will be referring to configurations such as (1) and (4a) as relations of *local* precedence, and to configurations such as (5) as *non-local* precedence. Based on the hierarchy in (3), we would then expect that, in local configurations, only the \(\alpha\) order should be possible, whereas both \(\alpha\) and \(\beta\) 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. 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 beta configuration is indeed less frequent than the alpha order, it is by all means attested. More careful analysis will however reveal that the beta order is only possible in those Slavic languages (Bulgarian, Macedonian and Bosnian-Serbo-Croatian) in which the polarity complementizer is a morphologically complex word resulting from the fusion of two independent clausal markers. To account for the specific subgroup in which the beta order is found, I will then 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 beta 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 alpha 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 structured as follows. In section II., I show that polarity questions are not islands for focalization. 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 author 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 beta as a grammatical possibility. In section IV., I establish a correlation between languages which display the beta order, and languages which have the option of realizing the polarity complementizer not simply as a free particle, but also an enclitic morpheme attaching to fronted XPs. Section V. illustrates the suggested derivation of polarity questions featuring a fronted focus in Macedonian, Bulgarian and Bosnian-Serbian-Croatian; in it, I argue in particular that the polarity complementizer in these three languages result 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, which I use to further refine my analysis of the irrealis particle da. Finally, in section VII., I briefly summarize the contents of the main proposal.

II. Polarity Questions Are Not Islands For Focalization

In his (2012) paper, Abels provides an elegant and economical way of accounting for precedence relations in the left periphery. Specifically, he argues that 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 may not precede the polarity complementizer because focalization is blocked by weak islands, and polarity questions, being questions, create weak islands.

---

3 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.
This take on word order restrictions introduces complete uniformity between local and non-local precedence relations: if the reason why a constituent in narrow focus may not precede POL has to do with the fact that the focus cannot escape the island created by this interrogative particle, we expect that configurations where a focus precedes POL will always be ungrammatical, regardless of whether they are local or non-local. This is because movement of the focus across POL will always result in ungrammaticality, regardless of what the landing site of the focus is. Indeed, this is one of the selling points of Abels’ analysis, and one of the reasons why this is a particularly economical way of accounting for restrictions on linear order: an identical set of principles – intervention effects – can be used to capture both local and non-local 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 (ITA)
   THIS refl. I-wonder if to-him(cl) they-wanted
dire (non qualcos’altro)
to-say (not something.else)

“It’s THIS that I wonder whether they wanted to tell him”

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. 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 two elements under discussion surface in an identical 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 (7-8) below, from Dutch, show. Dutch is an SOV, 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 POL is for such precedence relation to be non-local:
III. Focus and POL in Slavic Languages

Not all languages pattern with Italian and Dutch in only allowing the beta configuration non-locally: in some cases, a focus may also precede the polarity complementizer in local configurations. 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\(^4\) from Krapova’s original example:

\(^4\) The original observation that a focus may locally precede the Bulgarian polarity complementizer “dali” belongs to Ilyana Krapova. Example (10) differs from Krapova’s original example only with respect to the type of fronted focus: while this was simply contrastive in Krapova’s example, it is corrective (see Bianchi, Bocci & Cruschina 2016) in mine. Below is Krapova’s original example:

(i) Chudja se KNIGITE dali Ivan te vzeme (ili spisanijata) 
(1)wonder BOOKs-the whether Ivan will take (or journals-the) 
(Krapova 2002:109) 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 be fairly infrequent in other language groups as well; see for instance Cruschina (2016) on types of focus fronting in Romance languages.

7) 
(a) Ze vraagt zich af of MET JOHN je (NLD) 
*She wonders refl. vpt. if WITH JOHN I 
gesproken hebt (niet MET MARIE) 
spoken have (not WITH 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 (NLD) 
*WITH JOHN wonders she refl. vpt. if I 
gesproken hebt (niet MET MARIE) 
spoken have (not WITH MARY) 
“It’s WITH JOHN that she wonders whether I have spoken”
Example (10) illustrates how the Bulgarian polarity complementizer *dali* may either precede (cf. 10a) or follow (10b) a fronted constituent in narrow focus. Note that the fact that both the alpha and the beta order are acceptable in Bulgarian does not mean that these options are equally unmarked: at least my informants all seem to consistently prefer the order *dali* < FOC, and report that the opposite order feels either slightly more archaic.

Unsurprisingly, the non-local FOC < *dali* configuration is also acceptable. 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:

5 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 a problem.
Bulgarian is not unique in locally allowing for the order FOC < INT: this is also grammatical in Macedonian and, for at least some speakers, in Bosnian-Serbian-Croatian (henceforth, BSC). I report the relevant examples below. As the reader can see, these are modeled after Krapova’s example:

12)

(a) Не знам дали книгите ќе ги купи (MAC)

Не znam dali KNIGITE kje gi kupi
Not I-know if BOOKS-THE will them I-buy
(ne списанијата)
(ne spisanijata)
(not JOURNALS.THE)

(b) Не знам книгите дали ќе ги купи

Не znam KNIGITE dali kje gi kupi
Not I-know BOOKS.THE if will them I-buy
(ne списанијата)
(ne spisanijata)
(not JOURNALS.THE)

13)

(a) Pitam se da li ќе КNIJGE Ivan kupiti (BSC)

I-wonder refl. if will BOOKS Ivan buy
(ne ČASOPISE)
(not JOURNALS)

(b) %Pitam se КNIJGE da li ќе Ivan kupiti

%I-wonder refl. BOOKS if will Ivan buy
(ne ČASOPISE)
(not JOURNALS)

As it was already the case for Bulgarian, both Macedonian and BSC speakers tend to prefer the alpha order to the beta one, citing essentially the same reasons provided by Bulgarian speakers when judging the two word order configurations: the beta option sounds more archaic than the alpha configuration. Note that not all speakers of BSC found the order

---

6 Transliteration of Macedonian has been done according to the ISO9:1995 standard.
FOC < POL to be acceptable; I have therefore marked the corresponding example (13b) with an "%".

The beta 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 clitic če.

14)

(a) ??Sprášujem se, če bo KNJIGE Ivan (SLV)

??I-wonder refl., if will BOOKS Ivan

kupil, ne REVIJE!

buy, not JOURNALS!

(b) *Sprášujem se, KNJIGE če bo Ivan

*I-wonder refl., BOOKS if will Ivan

kupil, ne REVIJE!

buy, not JOURNALS!

The examples from Ukrainian, Polish, Czech and Slovak are reported below:

15)

(a) ?Хотів би я знати, чи книжки він (UKR)

?Xotiv by ya znaty, chy KNYZHKY vin

?Want cond. I to-know, if BOOKS he

купить, не журнали!

kupyt', nie ZHURNALY!

will-buy, not JOURNALS!

(b) *Хотів би я знати, книжки чи він

*Xotiv by ya znaty, KNYZHKY chy vin

*?Want cond. I to-know, BOOKS if he

---

7 A possible explanation for the overall lower acceptability of the β configuration in BSC will be offered in section IV.
купить, не журнали!
kupyt', nie ZHURNALY!
will-buy, not JOURNALS!

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

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

18)
(a) ?Zaujíma ma, či KNIHY Ivan (SLK)
?I-wonder refl., if BOOKS Ivan
vezme, nie ČASOPISY!
will-take, not JOURNALS!
(b) *Zaujíma ma, KNIHY či Ivan
*I-wonder refl., BOOKS if Ivan
vezme, nie ČASOPISY!
will-take, not JOURNALS!
We can thus identify two groups in the Slavic subfamily: on the one hand, we have languages such as Bulgarian, Macedonian and partially BSC, in which the beta order is a grammatical possibility even locally. Note that these all feature a morphologically 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 such as 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 below, from Slovene, shows how focus fronting across POL and to a higher left periphery is possible.

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

Ukrainian, Polish, Czech, Slovene and Slovak thus behave exactly like Dutch and Italian do: the β order is possible only if the focus has been fronted to a higher left periphery.

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 instances of “the books” appearing in a pre-dali/da li position ((b) examples in (10) through (13)) are in fact topics and not foci. On the one hand, (10-13) clearly depict a corrective speech act, and corrections
are arguably always focal, not topical\(^8\); moreover, if a topical reading of the pre-\textit{dali} constituent were responsible for the grammaticality of the beta pattern in BUL, MAC and BSC, we would expect the same to be possible in the other Slavic languages, contrary to fact. Still, this is an argument worth spending some time on, as it would indeed 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-\textit{dali} constituents would seem to find support in the fact that clitic doubling of the fronted “\textit{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 clearly regulated by mechanisms other than IS in Macedonian. In this language, the relevant criterion to determine whether a constituent should appear together with a coindexed clitic is rather whether such constituent is specific or not (Rudin et al. 1999; Kochovska 2010). In this respect, consider (20), which shows how a definite expression such as “the book/that

\(^8\) Regardless of what definition of “focus” one picks, the notion of “a topical correction” makes little sense. Consider the exchange below, where B’ contribution to the communicative exchange is clearly corrective in nature:

\begin{itemize}
  \item[i)] A: Mary likes Ricardo
  \item B: Mary likes KAREN (not RICARDO)!
\end{itemize}

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 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 by virtue of the fact that this is the only constituent representing new information. Concerning semantic definitions of “focus”, “KAREN” is also the only linguistic expression in (i) whose interpretation is clearly dependent on a set of alternative propositions (Rooth 1992, 1999).
book” must be resumed by a corresponding clitic. This is even though the constituent itself appears in situ:

20) (MAC)

Boris *(ja) прочитал книгата / овaa книга
Boris *(ja) procti*al knigata / ovaa kniga
Boris *(it) read book-the / that book

“Boris read the book/that book”

Now that we have established that the presence of a coindexed clitic in the Macedonian example in (12) is not necessarily an indicator of topicality, we can go back to reevaluating the argument according to which pre- *dali/ da li* expressions could be topics. As we cannot count on the presence or absence of clitic resumption to tell foci and topics apart, we must resort to other types of diagnostics. Something that we can do is to check 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 pre-*dali/ da li* position. 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*. Note also that this constituent is not clitic-resumed by a corresponding clitic like “*gi*” (=them), precisely because this is now a non-specific constituent:

21) (MAC)

(a) Се прашувам дали доволно книги
Se prašuvam dali DOVOLNO KNIGI
Refl. I-wonder if ENOUGH BOOKS
ке ку*пи*, (не доволно списаниja!)
ke kupi, (ne DOVOLNO SPISANija!)

---

9 A similar process seems to be taking place in Bulgarian: in colloquial Bulgarian at least, [+specific] 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 specific nominal expressions is mandatory, this process is still optional in Bulgarian.
will buy (not ENOUGH JOURNALS!)

“I wonder if s/he will buy ENOUGH BOOKS (not ENOUGH JOURNALS!)”

(b) Се прашувам довољно книги дали
Се прашувам DOVOLNO KNIGI dali
Refl. I-wonder ENOUGH BOOKS if
ке купи, (не довољно списаниja!)
ке купи, (не dovolno spisanjia!)
will buy (not ENOUGH JOURNALS!)

Example (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) (BSC)

(a) Pitam se da li je BAR KNJIGU Ivan
I-wonder refl da li aux AT.LEAST BOOK Ivan
kupio, ne BAR ČASOPIS
bought, not 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
I-wonder refl AT.LEAST BOOK da li aux Ivan
kupio, ne BAR ČASOPIS
bought, not AT.LEAST JOURNAL

23) (BUL)

(a) Чудя се дали достатъчно книги
Chudya se dali DOSTATACHNO KNIGI
I-wonder refl if ENOUGH BOOKS
той ще купи, (не достатъчно списания!)
toy shte kupi, (ne dostatachno spisaniya!)

“I wonder if she will buy ENOUGH BOOKS (not ENOUGH JOURNALS!)”

(b) Чудя се достатъчно книги
Chudya se DOSTATACHNO KNIGI
I-wonder refl ENOUGH BOOKS
dali той ще купи, (не достатъчно списания!)
We can conclude that constituents which are fronted to a pre-*dali*/*da li* position can indeed be foci in all the tree languages which make up group I.

### III.1 Overview

The table in (24) provides an overview of the Slavic languages reviewed so far, providing a bird-eye view of which languages locally allow for the $\beta$ order\(^{10}\), and which ones do not. The different languages are divided according to their respective subgroup:

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Language</th>
<th>Allows for $\beta$?</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Slavic</td>
<td>BSC</td>
<td>%YES</td>
</tr>
<tr>
<td>South Slavic</td>
<td>Bulgarian</td>
<td>YES</td>
</tr>
<tr>
<td>South Slavic</td>
<td>Macedonian</td>
<td>YES</td>
</tr>
<tr>
<td>South Slavic</td>
<td>Slovene</td>
<td>NO</td>
</tr>
<tr>
<td>West Slavic</td>
<td>Czech</td>
<td>NO</td>
</tr>
<tr>
<td>West Slavic</td>
<td>Slovak</td>
<td>NO</td>
</tr>
<tr>
<td>West Slavic</td>
<td>Polish</td>
<td>NO</td>
</tr>
<tr>
<td>East Slavic</td>
<td>Ukrainian</td>
<td>NO</td>
</tr>
</tbody>
</table>

Based on (24), 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 (6 if we include BSC) out of 8 languages not having this ordering configuration as a grammatical possibility. (24) also shows that the $\beta$ order is geographically circumscribed to the South Slavic area, although it is by no means a feature of all the languages in this particular

---

\(^{10}\) Unless otherwise stated, from now on all mentions to the $\alpha$ and the $\beta$ orders are to be understood as *local* instances of these configurations.
subgroup: Slovene rather patterns with Czech, Ukrainian, Slovak, Polish and Ukrainian in only displaying the α order.

One language notably missing from (24) is Russian. This is because Russian has no free polarity morpheme: in this language, when POL is lexicalized, it takes the form of the enclitic morpheme -ли (=-li)\textsuperscript{11}. Russian -li must either encliticize onto a finite verb (25), or onto a narrow focus if such constituent is present (26):

25) Я не знаю, работает ли он на заводе (RUS)
Ja ne znaju, rabotaet li on na zavode
I don’t know, works -li he at factory.PREP

26) Книгу ли он читает? (RUS)
Knigu li on citaet?
BOOK.ACC -li he reads?
‘Is it THE BOOK that he is reading?’

In Russian, the relative order of fronted narrow foci and the morpheme expressing POL is thus always fixed in the β configuration as a result of the enclitic nature of POL itself.

Another East Slavic language missing from (24) is Belarusian. Belarusian does not feature in (24) for the same reason Russian does not: this language has no free morpheme to express POL. Exactly like Russian –li, Belarusian цi- (= ci-) takes as its prosodic host either the fronted finite verb, or a constituent in narrow focus. Unlike –li, however, ci- is proclitic on its phonological host:

27) Цікава, ці чытала яна книгу (BEL)
Cikava, ci chytala iana knigu
I-wonder, ci- read she book.ACC

‘I wonder if she read the book’

\textsuperscript{11} Note however that, 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.
With Belarusian, we thus observe a situation opposite of that observed in 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”.

IV. Complex Complementizers

In the previous section, we saw that Slavic languages may be divided in two groups with respect to the possibility of locally fronting a focus to a pre-POL position. 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 Italian and Dutch) a locally fronted focus may only follow the polarity complementizer.

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 if this is present. 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. If this is the case –as already seen– 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\(^\text{12}\). That POL may 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, although there is a general tendency, observed in all three languages, to favor the use of dali/ da li over -li in embedded questions. The use of the –li strategy is illustrated in (29-30) for matrix polarity questions, and in (31) for embedded ones:

\[
\begin{align*}
29) & \text{ ИВАН ли рисува всеки ден?} & \text{(BUL)} \\
& \text{IVAN li risuva vseki den?} \\
& \text{“Is it IVAN who paints every day?”} \\
& \text{(Dukova-Zheleva, 2010: 177)}
\end{align*}
\]

\[
\begin{align*}
30) & \text{ Рисува ли Иван всеки ден?} & \text{(BUL)} \\
& \text{Risuva li Ivan vseki den?} \\
& \text{“Does Ivan draw every day?”} \\
& \text{(Dukova-Zheleva, 2010: 170)}
\end{align*}
\]

\[
\begin{align*}
31) & \text{ Pitam se studira li Marko medicinu} & \text{(BSC)} \\
& \text{Wonder-I refl. studies -li Marko medicine}
\end{align*}
\]

\(^{12}\) 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) 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 Engelund’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.
“I wonder if Marko studies medicine”

Note that dalı may equally be employed to form matrix polarity questions as well, a phenomenon that we will appreciate more in detail in the coming sections.

We may conclude that dalı and li are functionally equivalent, in the sense that both elements type a given proposition as being a polarity question. These elements being functionally equivalent does not mean that the two types of questions are also pragmatically equivalent, however. In both Macedonian and Bulgarian, for instance, whereas dalı 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. Differently put, whereas an embedded dalı 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.

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:

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

Why would a generalization (32) 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, which is that those languages which locally allow for β are the same languages which get the β order ‘for free’ whenever POL is realized as an enclitic. This is because, as we just saw, -li must encliticize onto a constituent in narrow focus if this is present. When this is the case, the β configuration obtains automatically as a result of the enclitic rather than proclitic nature of -li itself.

13 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.
In its current formulation, (32) is however 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 for Czech is illustrated in (33), which I take from Toman (1996):

\[33) \text{Nevíme, mají -li dnes medovinu} \quad (\text{CZ})\]

\[\text{Not-we-know, they-have -li today honey}\]

'We don’t know whether they have honey today'

Toman (1996:508)

Clearly, some refinement to (32) is needed.

What seems to be the relevant factor in capturing the group I/ group II asymmetry is not simply whether a given language has an enclitic realization of POL, but also the specific type of constituent such morpheme may attach to. Specifically, whereas all languages which feature the morpheme –li as an enclitic polarity particle can have this encliticize onto a fronted verb, only some languages allow for it to encliticize onto an XP. Consider in particular the table below, which I take from Schwabe (2004:10)\(^{14}\); (34) details grammatical enclisis sites for the –li morpheme in the various Slavic languages which feature this morpheme as part of their inventory. In (34), a full circle indicates productive usage, whereas an empty circle indicates archaic usage.

\[34) \]

<table>
<thead>
<tr>
<th></th>
<th>V-li</th>
<th>XP-li</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Matrix</td>
<td>Subord.</td>
</tr>
<tr>
<td>SS</td>
<td>BSC</td>
<td>●</td>
</tr>
<tr>
<td>BUL</td>
<td>●</td>
<td>●</td>
</tr>
<tr>
<td>MAC</td>
<td>●</td>
<td>●</td>
</tr>
</tbody>
</table>

\(^{14}\) Note that Schwabe’s original table is more detailed than the one I have provided in (34), 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.
(34) shows us that both Bulgarian and Macedonian 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 maximal projection, but unlike what observed in Macedonian and Bulgarian, this option is perceived as more archaic than the verbal enclisis option. According to Rudin et al. (1999) (see also references therein), XP-li in modern BSC also tends to be limited to single-words XPs and is not accepted by all speakers.

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 \(-li\)\(^\text{15}\); this is also supported by Toman (1996), who reports that \(-li\) structures are only limited to V-li. I have thus marked the corresponding cell with an asterisk. Conversely, at least some of my BSC informants accepted \(-li\) as enclitic onto fronted XPs; these were those same speakers who accepted the \(\beta\) order.

In BUL, MAC and, for at least some speakers, in BSC, the enclitic POL morpheme is thus a flexible element, capable of encliticizing either onto the finite verb, or onto a fronted XP. In Czech, on the other hand, \(-li\) may only attach to the finite verb. This leads us to revise (32) as follows:

\(^\text{15}\) Of course, this does not mean that no speaker of Czech ever admits the possibility of encliticizing \(-li\) onto an XP; it simply means that none of my informants did.
35) Only those languages (i) which have the option of realizing the polarity complementizer as the enclitic morpheme –li (ii) where such morpheme may encliticize onto an XP, have the option of having a fronted focus locally precede POL.

Why would a revision of (32) in the form of (35) be more apt to capture the group I/ group II distinction? Recall that we are interested in whether a given language has the possibility of having the interrogative complementizer follow a focalized XP. The revised generalization in (35) fleshes this out by specifically establishing a connection between languages which display the β order, and languages in which –li may also encliticize onto a focalized XP rather than simply onto the fronted verb.

V. Da Meets –li, Polarity Ensues

Now that we possess a working generalization capturing exactly what Slavic languages allow for beta locally, we want to be able to incorporate such a generalization into a theory accounting for why this might be possible in the relevant language group.

As already briefly remarked in section III., all languages which allow for β display a POL complementizer which has an identical form: 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. In this paper, I would like to extend the diachronic incorporation analysis of dali/da li to the synchronic dimension, and argue that present-day dali/da li still results from the incorporation through movement of the irrealis clausal marker da together with the enclitic polarity morpheme –li.

Before we discuss the proposed derivation leading to the emergence of the complex dali/ da li form, 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\(^1\), the non-factive marker in BSC, MAC and BUL is its own morpheme. In all the 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* mark an independently true statement, *da* carries no such implication. This is exemplified in (36) for Bulgarian, which I take from Rudin (1986). Note in particular how *če* must be used if it is the case that the speaker was indeed hungry:

36) (BUL)

(a) Не усещах че съм гладен
    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)

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.

\(^1\) Obtained by incorporating the factive complementizer together with the conditional marker, see Russian чтобы (*čtoby*), resulting from the union of the declarative complementizer *eto* together with the conditional particle *by*.
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:

37) (BSC)

(a) Ana misli da Marko spava  
Ana thinks that Marko sleeps  
“Ana thinks that Marko is sleeping”

(b) Ana smeta što Marko stalno spava  
Ana bothers that Marko always sleeps  
“It bothers Ana that Marko is always sleeping”

(Mihaliček 2012: 114)

A second function performed by da pertains to the expression of what would be an infinitival structure in languages like English. Neither Macedonian nor Bulgarian possess a system of infinitives like English does\(^\text{17}\); in these two languages, what would be an infinitive structure in English may be realized by having da procliticize onto the embedded verb. This is illustrated for Macedonian with the contrast between (38a) and (38b). Note in particular that the verb onto which da procliticizes still retains person and number features:

38) (MAC)

(a) Заборавив дека седам во канцеларија  
I-forgot that I-sit in office  
‘I forgot that I was sitting in an office.’

(b) Прозорецот заборавив да го затворам  
The-window I-forgot that it I-close  
‘I forgot to close the window.’

(Lindstedt 2010:416)

\(^{17}\) 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 the “infinitive”. 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.
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 (38b) -where *da* introduces as embedded clause- as *subjunctive da* structures.\(^{18}\)

A third function relates to the expression of modality. In particular, when *da* appears sentence-initially, it signals the presence of a desire or a weak command (39a). In this case, *da* performs a function similar to that performed by the subjunctive mood in matrix sentences in languages like Italian (39b):

\[\text{(39)}\]
\[
\begin{align*}
\text{(a)} & \quad \text{да живее Европейският съюз} \quad \text{(BUL)} \\
& \quad Da zhivee Evropejskiyat sajuz \\
& \quad Da \text{ lives European, the union} \\
& \quad \text{“Long live the European Union!”} \\
\text{(b)} & \quad \text{Che tu possa vivere a lungo!} \quad \text{(ITA)} \\
& \quad That you may(subv) to-live a long \\
& \quad \text{“May you live a long life!”}
\end{align*}
\]

In matrix polarity questions, on the other hand, *da* signals the presence of deontic modality (40):

\[\text{(40)}\]
\[
\begin{align*}
\text{(BUL)} & \quad \text{да се върне ли довечера?} \\
& \quad Da se varne li dovechera? \\
& \quad Da refl. (s)he-returns li tonight? \\
& \quad \text{“Should (s)he return tonight?”} \\
\text{(Rudin 1986: 118)}
\end{align*}
\]

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

\[\text{(41)}\]
\[
\begin{align*}
\text{(MAC)} & \quad \text{да го добиевме твоето писмо,} \\
& \quad Da go dobievme tvoeto pismo, \\
& \quad Da it we-had-received your letter, \\
& \quad \text{“If we had received your letter, I would have come”}
\end{align*}
\]

\(^{18}\) See however Lindstedt (2010) on why using the term ‘subjunctive’ may be problematic.
Summing up, *da* appears to perform a multitude of different functions in the 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 (42). In (42a), the embedded subject occurs in a post-verbal position; if locally fronted through focalization or topicalization, this must crucially precede *da* (42b vs. 42c).

42)  
(a) Искам да пеят децата  
    Iskam da peyat detsata  
    *I want the children to sing*  
(b) Искам детсата да пеят  
(c) *Искам da детсата пеят*  

(Rudin 1983:4)

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 subject of the embedded clause, which must either precede *da* (43a) or follow the embedded verb (43c).

43)  
(a) Искам тя да дойде  
    Iskam tya da doyde  
    *I want her to come*  
(b) *Искам da tya doyde*  
(c) Искам da doyde tya  

(Lindstedt 2010:416)

Concerning the status of *da*, this morpheme 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 complementizer status.
VI. The derivation

(44) below illustrates the proposed derivation for sentences featuring *dali* and a fronted narrow focus. (44) 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, (44) is inevitably underspecified with respect to some features\(^{19}\). The derivation in (44) 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 POL 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 assuming that the ultimate landing site of POL will be higher than the position in which foci are externally merged; other than the order FOC < POL being cross-linguistically rarer than POL < FOC, this is something for which I have not yet found independent justification\(^{20}\).

In (44), *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. The resulting complex form *dali/dali* is then moved to the head of IntP, the position in which polarity complementizers are hosted according to Rizzi (2001):

\[
(44)
\]

\(^{19}\) Some features that the derivation in (44) 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 Wackernagel second position.
- The fact that *-li* in Macedonian may split a focused, where Bulgarian *-li* follows the entire DP even in those cases where only part of such constituent is focalized.

\(^{20}\) See however Bianchi & Cruschina (2016) for a sketch of a possible semantic explanation of the higher position of POL wrt FOC.
Let us discuss the details of (44), starting with the proposed external merge site for *da*.

In (42-43), we saw that *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 to originate in *Fin*, the lowest functional projection capable of hosting complementizers in Rizzi’s hierarchy (see again (3)). Note that positing that *da* is in Fin would be required independently of what its distribution might suggest because of the specific framework we are adopting: according to Rizzi (1997), Fin is the functional projection which is dedicated to the hosting of markers of [-realis] subordination.

As already mentioned in the preceding section, South-Slavic *da* is not a free morpheme but a clitic. Unlike *-li*, *da* is proclitic: it must linearly precede its phonological host. In most cases, it will be a verbal head which provides *da* with a site onto which to procliticize: this will result in the verb-adjacency pattern we observed in (42-43). Verbs are however 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* 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.

The process of double-cliticization (i.e., –*li* encliticizing onto *da*, and *da* procliticizing onto –*li*) I am suggesting to be underlying *dali* is by no means unheard of in Slavic languages. Consider the following examples, which illustrate the use of the forms *ne li/ neli*, from Macedonian and Bulgarian respectively. In Bulgarian and Macedonian, the negation marker *ne* is proclitic onto the verb: nothing with the exception of other clitics\(^21\) may appear in between *ne* and the verb this procliticizes onto. In (45), we see that it is possible for –*li* to encliticize not simply onto a finite verb or onto a fronted, focussed XP, but also onto the negative marker itself; this results in the creation of a negative polarity-question interpretation.

45) не ли сакаш да одиш?
   Ne  li  sakaš  da  odiš?
   *Not  -li  you-want  da  you-go*
   “Don’t you want to go?”

(Rudin et al. 1999: 556)

(45) is thus evidence that the element functioning as a host for a clitic may itself be a clitic. (46) illustrates the distribution of –*li* with respect to pronominal clitics in negative polarity questions. In particular, we see that the relative position of –*li* within the clitic cluster is different in the two languages. Bulgarian *ne* has peculiar phonological properties (Hauge 1999) which differentiates it from Macedonian *ne*: it causes whatever element follows to be stressed, even in those cases when this element is another clitic. This is illustrated in (46a), where the pronominal clitic *go* appears right after *ne*. As stress is assigned to *go*, the string “*ne go*” forms a phonological unit that –*li* may encliticize onto, resulting in the word order we observe in (46a): –*li* appears in third position, after both proclitics (Izvorski et al. 1997).

46)

(a) не го ли видя?
   Ne  go  -li  vidya?
   *Not  him  -li  s(he).saw*
   “Didn’t she see him?”

(Rudin et al. 1999: 556)

\(^21\) The exact relative order of clitics inside the clitic cluster is the following:

(i) Negation > Auxiliary > Dative > Accusative > VERB
Macedonian `ne` does not equally result in the following clitic to be stressed, but is rather stressed itself. Even when a pronominal clitic is present, `-li` will then still encliticize onto `ne` (46b):

(b) Не ли го виде?  
(Not `li` him (s)he.saw
“Didn’t she see him?”

Finally, consider (47), which displays the complex “neli” form (written as a single word), which is used as a negative question tag in Macedonian. The difference between (46b) and (47) does not simply relate to spelling but also pertains to stress assignment: whereas in (47) the two morphemes are phonologically distinct (with syllabic stress being assigned to the negative particle, “NE li”), the two morphemes form a single phonological word (“neLI”) in the case of (47).

47) Нели сме му го дале?  
(Neli we.are to.him it given
“We gave it to him, didn’t we?”

The fact that a form like “neli” exists shows us that there is a tendency for –`li` to form complex words by encliticizing onto other clitics, which is exactly the derivation I am suggesting for `dali`.

Quite interestingly, there is some controversy in the literature (see in particular Rudin et al. 1999 and references therein) concerning whether “neli” the negative tag, and Macedonian “ne li” as found in (46b), should actually be considered as two different forms. According to Tomić (1996), for instance, the appropriate translation of examples like (46b) should always be that of a normal, positive polarity question plus accompanying negative tag (i.e, “you want to go, don’t you?”). Rudin et al. (1999), on the other hand, report that at least some of their informants accept a non-tag interpretation of questions like (46b). According to the authors, this is likely due to internal differences in stress assignment patterns in Macedonia. In particular, some speakers of Macedonian would process “ne” the way Bulgarian speakers do, namely as a particle capable of assigning stress to the element to its
immediate right. If this happens, stress is assigned to \textit{–li}, and \textit{“ne+li”} is processed as a single phonological word, generating \textit{“neLI”} and hence a negative-tag interpretation. If \textit{ne} is processed “the Macedonian way”, on the other hand, no syllabic stress will be assigned to the element to its immediate right but rather it will be \textit{ne} which receives stress. This allows for the possibility of having strings of the type of \textit{“NE li”}. If this is an option, it follows that there will indeed be a distinction between \textit{“NE li”} and \textit{“neLI”}, which in turn will generate the possibility of having two different interpretations of sentences like (46b) and (47): a negative-polarity question one, and a negative-tag question one. The fact that the correct interpretation of \textit{“ne+li”} strings in questions is relatively controversial is evidence suggesting that complex strings resulting from the union of \textit{–li} together with another proclitic element are still fairly transparent in their composition; in turn, this substantiates our analysis of \textit{“dali”} as still resulting from the movement plus incorporation of two separate clausal markers.

We have now accounted for why \textit{da} would move up to combine with -\textit{li}. The derivation is however not complete: if the complex morphological word \textit{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 POL to always follow constituents in focus in BSC, MAC and BUL. Differently put, we would expect the \textit{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 \textit{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 \textit{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 polarity complementizers according to Rizzi (2001). We have now derived the order \textit{POL < FOC}, and the derivation is finally complete.

How does the derivation in (44) capture the fact that group I languages may display both the \textit{α} and the \textit{β} order? Consider again (44). There is a point in the derivation in which the complex morpheme \textit{dali} is found in a position \textit{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 \textit{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 exactly what copy in their movement derivation one may spell out. In particular, they postulate that all mirative and corrective foci must always (whether covertly or overtly) move to the specifier of a left-Peripheral Focus projection, as it is only in this position that the specific conversational implicature
associated with these two types of foci may be licensed. Nothing however excludes that a
copy other than the one representing the head of the movement chain (i.e. the one in
Spec,FocP) may be given overt spell out instead. According to Bianchi, Bocci and Cruschina,
this captures the fact that, in structures where, say, a corrective focus is present (48), such
focus may either front or remain in situ, with no apparent consequences on the meaning or
the truth conditions of the sentence:

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
POL. 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 in the head of the
movement chain (=the head of the Int projection), or the one located in the head of the
Focus projection (=where da and -li are combined). Note that it would not be possible to
give overt spell out to any other copy of da, in that this morpheme, as already mentioned,
is not free and thus needs a phonological host onto which to attach.

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 found in the specific group of
languages we have identified. Recall that, according to the analysis we have developed, the
reason why group I languages display the possibility of having FOC locally precede POL is
because, at some point in the derivation of a focused polarity question, POL 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 POL in these languages: in group I
languages, POL results from the combination of an enclitic morpheme specified as [+ focus]
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 exhibited by group II languages are not
as 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. The crucial difference between Czech jestli and Macedonian, Bulgarian and BSC
dali/da li is that jestli does not arise from the union of two subordination markers, but
rather from the union of the verbal form jest (former 3rd person sg of být = to be, no longer
in use) and –li (Tabakowska 1997). 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 lexicalized -and hence no longer
transparent- type of polarity complementizer.

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 POL is
generated directly in Int, at no point in the derivation of an embedded polarity question
with a fronted focus will POL be found in a position preceding the focalized constituent;
this is because POL 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 POL 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 POL is merged. Indeed, this is precisely what we observe in group II languages:
β is only possible as part of a non-local configuration.

The analysis I have just presented in very similar in spirit to the analysis presented in
Krapova (2002) to account for the fact that Bulgarian exhibits both the β and the α order
(see again footnote 3). Rather than accounting for the observed optionality by suggesting
the existence of a second focus position -a logical, although problematic, possibility-,
Krapova suggests in particular that POL is generated in a position lower than Focus, and
may optionally raise to a position above it. She then postulates that optional movement of
POL accounts for languages like Bulgarian, which displays optionality between alpha and
beta, whereas mandatory movement of POL to this pre-Focus position accounts for
languages like Italian, where only the alpha configuration is observed. My account patterns
with Krapova’s in accounting for the existence of the β order by assuming movement of
POL, but unlike hers, takes POL to be moved rather than base-generated only in a specific
subset of languages.
VI.I Where Does -Li Go?

In (44), -li is the head of the Foc(us) projection. Assuming that -li is connected to the expression of focus is 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 (49) below.

49) глядя ли его?       (BUL)
    Gleda li go?
    Looks -li it(cl)?
    ‘Is (s)he looking at it?’

-Li questions in which it is the finite verb which fronts to the left periphery have traditionally been referred to as “neutral”, in a bid to distinguish them from questions where it is an XP which fronts, something which –as we saw– results in the appearance of a narrow-focus interpretation of the fronted element. In fact, yes/no questions with the structure of (49) can be equally –context permitting– interpreted as being narrowly focalized: a possible interpretation of (49) 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:

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

Note also that even 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. It is also possible to argue that -li is additionally specified with a [+ interrogative] feature, since -li structures are not simply narrow focus structures, but they are also questions as well. This 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 morpheme in polarity interrogatives typing the

---

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°.
clause as being a question. In fact, however, it is also possible 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. In this respect, consider the case of Burkinabè Bambara, which equally possesses a dedicated morpheme to mark the presence of a narrow focus. This morpheme, lo, is crucially very different from the morpheme used to mark a proposition as being a polarity question, wa (see [author, in preparation]). This is illustrated in (51a-b) below. Quite 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.

51) (BAM23)

(a) Seydou ka dji mi (wa)?
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’

Evidence against the superimposing of the [+focus] and [+interrogative] features in our analysis of -li comes from the following Macedonian example, which I take from Friedman (2018)24. (52) illustrates how -li may also have a purely emphatic, non-interrogative interpretation: in (52), -li is used in conjunction with the expression of a dubitative structure.

52) Kako da ne, toj li kje ti ja napravi kolata...
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 (52) is a bit of a “Serbism” stylistically speaking, it is perfectly acceptable in Macedonian.

In this paper, I will then remain agnostic as to whether –li also spells out a [+interrogative] feature, and simply assume that it minimally spells out a [+focus] feature.

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

23 As spoken in Burkina Faso. Bambarà is also spoken in Mali (where it is the national language), Côte d’Ivoire, Gambia, Senegal, Guinea, Sierra Leone and Ghana.
24 Thanks to Victor Friedman for suggesting this example to me.
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*.

VII. “*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 will see that
one such particular combination, “dali da”, will warrant 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 (53). (53) features a *wh*-word, already
typing the clause as being a question, as well as the polarity complementizer *če* and the
irrealis marker *da*:

\[ \text{53) Kdo če da pride?} \]  
\[ \text{Who if that comes?} \]  
\[ \text{“Who is said to be coming?”} \]  

(Hladnik 2010:15)

Likewise, both –*li* and *da* may be found in combination with other elements already typing
a clause as being interrogative or irrealis, in all of 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.\(^{25,26}\)

---

\(^{25}\) 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 VI.I). If –*li* forces a focus interpretation of the constituent which
is fronted to a pre-*li* position, it follows that the interpretation of (54) is roughly equivalent to English
“but WHO has time for that(?)”. Here, emphatic stress on the interrogative constituent can be used to
convey, for instance, that the existential implicature generally associated with *wh*-elements –in this case,
that some x must exist such that x has the time to do ‘that’- must be canceled, as no such x presumably
exists.
54) Но кой ли има време за това?  

*But who -li has time for that?*  

(From *The Bulgarian National Corpus*)

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

In its current formulation, the derivation in (44) has the advantage of capturing the ungrammaticality of one such combination, namely the fact that a sequence of the type of "*dali li" is never possible in any of the three languages under discussion. That "*dali li" should never be possible 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 typing the clause as a yes/no interrogative: in both cases, this is generated in an identical projection. This automatically accounts for the fact that "*dali/ da li" and "-li" in isolation are in complementary distribution with each other. In this respect, (44) 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 (44) 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 (55) below:

55)  

(Dukova-Zheleva 2010:55)

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

If (55) were correct, however, we would expect sequences of the type of “*dali li” to be possible, 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 are at least functionally equivalent in that they both type the clause as being a polarity question. This naturally follows from (44) 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 (55)\textsuperscript{29}.

If the derivation in (44) 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 in fact very frequent. An example of the relevant structure is provided in (56) for Bulgarian, and in (57) for BSC.

(56) is an embedded polarity question; in this case, the addition of (a second) da results in what is usually translated into English by resorting to an infinitival structure\textsuperscript{30}:

\textsuperscript{29} A possible combination of clausal markers that the derivation in (44) would appear to correctly rule out is “*da dali”, on the grounds that (i) there is a single head capable of lexicalizing da in the left periphery, not two, and (ii) no copy of da is ever found in a position higher than the one in which dali is created. This is however a false positive: we expect “*da dali” to be ruled out independently because da is a proclitic element requiring either a verb or a clitic to attach onto.

\textsuperscript{30} Recall from section V that neither Bulgarian nor Macedonian possesses a system of infinitives: the particle da is used instead. The verb which would be infinitival in languages like English is then conjugated in the present tense form. Below are a series of da-structures to illustrate the mechanism.

(i) исках да дойда с вас (BUL)
Исках da doyda s vas
I-wanted da I-come with you
“\textquotesingle I wanted to come with you”

(ii) исках да дойдеш с мене (BUL)
Исках da doydesh s mene
I-wanted da you-come with me
“I wanted you to come with me”

As we can see from (i-ii), a subjunctive da structure is still not quite identical to what would be an infinitival structure in English. Although the tense of the verb which is selected by da is fixed (always present tense, regardless of the tense of the matrix verb), the verb person features. Also note that, when the pronominal subject is realized (both Bulgarian and Macedonian are pro-drop), this must be in the nominative case:

(iii) искам тя да дойде с мене (BUL)
Искам тя da doyde s mene
I-want she(nom) da (s)he-comes with me

“I want her to come with me”
Theon unintentionally wondered whether to say prayers.

From The Bulgarian National Corpus

In (57), a matrix yes/no interrogative, the addition of the second da results in the expression of deontic modality: specifically, the question shifts from being a simple polarity question (“Is Vesna reading this book?”) to an inquiry on whether Vesna should read the book.

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 polarity questions where only dali is present, and questions where a second da is realized as well. In section V, I argued that the “da” we observe in dali/ da li originates in a low left-peripheral position, only to raise up to combine with the focus marker –li. To capture the existence of “dali da” structures, one could then argue that the second da is nothing but the spell-out of the external merge site of da, i.e., the spelling out of the foot of the movement chain of da. If this analysis were correct, however, 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 dramatically.

The key to understanding how to account for “dali da” strings lies, I believe, in taking into account the remarkable versatility of the marker 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, counterfactuals and subjunctive structures. To capture the extremely versatile nature of da, as well as the existence of strings of the type of “dali da”, I will be adapting a proposal originally presented in Todorović et al. (2017) to account for the role and distribution of da in BSC. Todorović et al. 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 argued that this clausal marker is employed in MAC and BUL to realize the equivalent of what would be an infinitival structure in languages like English. It however makes more sense when describing BSC in that, unlike BUL and MAC, at least some varieties\(^{31}\) 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 (58) below:

\[
\begin{align*}
\textbf{58)} & \quad \text{(BSC)} \\
& \quad \text{(i) Odlučila} \quad \text{sam} \quad \text{da} \quad \text{prevodim} \quad \text{pesmu} \\
& \quad \text{Decided} \quad \text{aux.1.SG} \quad \text{da} \quad \text{translate.1SG.PRES} \quad \text{poem} \\
& \quad \text{(ii) Odlučila} \quad \text{sam} \quad \text{prevedem} \quad \text{pesmu} \\
& \quad \text{Decided} \quad \text{aux.1.SG} \quad \text{translate.INF} \quad \text{poem} \\
& \quad \text{“I decided to translate the poem”} \\
& \quad \text{\textbf{(Todorović et al. 2017: 2)}}
\end{align*}
\]

(58i) has the same meaning of (58ii), 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 et al. 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 this features. The authors also suggest that *da* is capable of spelling out different functional heads (C, T or little v), depending on the predicate selecting for the verb whose finiteness *da* goes to spell out. Following Wurmbrand (2001, 2014, 2015), and on the basis of phenomena like clitic climbing\(^{32}\), Todorović et al. 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:

\[
\begin{align*}
& \quad \text{(i) She decided to eat tomorrow} \\
& \quad \text{(ii) She tried to eat *(tomorrow)}
\end{align*}
\]

\(^{31}\) See again footnote 17.

\(^{32}\) On the assumption that CPs are barriers for such a phenomenon, see Wurmbrand (2015).
In BSC, both (59i-ii) could be expressed with a *da* structure, but crucially the *da* morpheme would then be spelling out two different clausal layers: the TP in (59i), and the vP in the case of (59ii).

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

60) **Kaže da će da dođe.**

*Says da will.3.SG da come.3.SG.PRES*

‘He says he will come.’

(Todorović et al. 2017, from Sočanac 2011: 55)

Todorović et al. analyze (60) by suggesting that, whereas the higher *da* spells out C, the second, lower *da* spells out v. This is illustrated in (61):

61) 

![Diagram](image)

(Todorović et al. 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. Moreover, even in BSC, assuming that all that *da* does is overtly spelling out the finiteness of any verb on which this morpheme may be proclitic seems slightly reductive. This is particularly so given that, even in BSC, the addition of *da* in a matrix yes/no *dali* question clearly results in the emergence of deontic modality, as already illustrated in (57), which I repeat below. Compare in particular (57) with its *da*-less counterpart in (62): as the reader can appreciate, the meaning is indeed quite different and arguably not simply reducible to the presence or lack of an overt [+finite] feature.

57) **Da li da Vesna pročita ovu knjigu?**

*Da li da Vesna reads this book*

‘Should Vesna read this book?’
In this paper, I will be following Todorović et al. (2017) in assuming that da occur more than once within the same clause, and that when it does, the different das spell out different clausal layers. Unlike Todorović et al., 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. Labels other than the [−realis] one could in principle be envisaged to capture the different functions da performs in the languages under discussion, but anything more specific than [−realis] would be at the price of losing sight of the extreme versatility of this clausal marker. Rather than coming up with more specific labels, I will then be assuming that, whereas da always spells out the presence of [−realis], the exact compositional import of this [−realis] feature hinges on whether da is selected or not by a higher verb, and on what functional projection da spells out. For instance, whereas unselected da expresses deontic modality or optativity (see (57)), a da which is selected by a higher verb is involved in marking (non-factual) subordination (see for instance (36) and (37)). Consider also the following example, where it is clear that the two das perform two different functions: in particular, only the first da (unselected by any higher predicate) is involved in the expression of optativity.

63) *Da* otdem *da* pijem po edna bira!  
*Da* we.go *da* we.drink on one beer!  
“Let’s go for a beer!”
64) Да почукаям ли?
   Da pochukam li?
   “Should I knock?”

We can assume the following derivation for (64): the verb *pochuka* is externally merged in
v, and *da* is merged as the head of T, marking the fact that the predicate is to be associated
with deontic modality. 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:


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

66) Да злорадстваш ли дойде?
   Da zloradstvash li doyde?
   “Have you come to gloat?”

(64) and (66) are at least partially string-identical, but their derivation is crucially different.
In particular, in (66), what is fronted to a pre- *li* position is not the closest available verb,
but its subordinate predicate:


---

31 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?
   “Do you want me to die?”

(ii) Да умра ли искаш?
    Da umra li iskash?
    “Do you want me to die?”

From *The Bulgarian National Corpus*
The fact that the two das in (64) and (64) have different interpretations follows given their respective extraction sites: whereas the fronted da in (66) was extracted from an embedded clause, the da in (64) originates from a matrix one. A deontic interpretation is thus only available for (64).

Now that we are familiar with the mechanics of 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 et al. (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 et al.) 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. Consider strings of the type of “dali da”: if da always spells out [-realis], what reason would there be to also do so at the level of the CP, when the lower da already performs the same function for that specific clause?

The crucial idea I will be adopting is that spelling out [-realis] at the level of the CP results in an effect which is substantially different from spelling out [-realis] in a lower projection, which justifies spelling out two separate da morphemes. Earlier on I have shown that the specific import of the [-realis] feature spelled out by da is dependent on whether this is selected or not by a higher predicate. Assuming that the nature of this [-realis] feature that da goes to spell out is also dependent on the specific clausal layer this is merged in (C or T/v) is simply an additional step in that same direction.

Before we move on with our analysis, it is useful to consider again the differences between dali questions on the one hand, and –li questions on the other. In section IV., we saw in particular 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-clausal-layer da, logic then 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?
Mapping the diachrony of how exactly [-realis] came to be connected to the expression of subordination in Slavic is beyond the scope of this paper, which is mainly concerned with accounting for the existence of cross-linguistic variation in the relative order of foci and interrogative complementizers. Nevertheless, we may attempt to provide a sketch of an explanation at least.

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. 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 (68). In this example, the subordinate clause rakharana is not in the subjunctive mood. According to Bickel (1993), since it is not subjunctive, the interpretation of this clause is actually ambiguous, the possible interpretive differences relating to the degree of integration of said dependent predicate with the matrix:

68) Rak-khar-a-na hab-he i?

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 (68a) as an instance of “illocution attraction”: the dependent predicate remains in the scope of the interrogative illocutionary force associated with the matrix predicate. In (68b), on the other hand, the dependent predicate retains independence from the illocutionary force of the main predicate.

If the use of subjunctive mood and that of [-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 (68). Specifically, we may suggest that the [-realis] marker da merged in CP acts as a barrier for illocutionary force in ways similar to what the subjunctive mood does in Belhare. The idea is the following: in an embedded dali question, the additional 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

---

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.
other hand, since the extra *da* is absent, 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 is inconsequential in that 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.

**VIII. Conclusions**

The goal of this article was to investigate the relative distribution of two types of left-peripheral constituents: fronted constituents in narrow focus, and polarity complementizers (POL). I have focused in particular 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 POL) is locally never possible.

Contra Abels (2012), I have argued that the ungrammaticality of the beta configuration in those languages where this is not possible 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 POL would be equally ungrammatical, whereas this is never the case.

Not all languages disallow foci from locally preceding POL. In the Slavic subgroup, for example, two (three if we count Bosnian-Serbian-Croatian, for which acceptance of beta seems considerably less widespread) out of the eight languages which possess a free morpheme to realize POL allow for beta 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 beta may be possible locally in these specific languages, I have argued, is no
accident, but rather a consequence of the morphologically complex nature of POL in the
three languages under discussion. In MAC, BUL and BSC, POL is spelled out as dali/ da
li. Crucially, both of the two morphological building blocks which go to form this word,
namely -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 POL 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 may be spelling out different clausal layers and which is
proclitic in nature. The second element is the structurally higher [+focus] marker –li, an
enclitic morpheme. This incorporation process not only provides a phonological host for
both clitics, it also correctly types the polarity complementizer as introducing a [-realis],
focus-sensitive environment. In all the languages where the alpha order is the only option
locally, on the other hand, the morpheme which lexicalizes POL is either morphologically
not complex or no longer transparent, and is thus likely not to result from the incorporation
of distinct clausal markers.

Crucially, because a movement-plus-incorporation analysis underlies the derivation of POL
in Macedonian, Bulgarian and BSC, a copy of POL itself is present in a position lower than
the one in which the focus is internally merged when fronted. This generates optionality
with respect to exactly what copy of POL may be given overt spell-out: when the copy
which is pronounced is the one found where the incorporation of da with -li takes place, the
beta order ensues. When the head of the movement chain of POL is pronounced, on the
other hand, the alpha order is obtained.

IX. References

229-254.
Bianchi, V.; Bocci, G.; Cruschina, S. (2016). Focus fronting, unexpectedness, and evaluative
implicatures. Semantics and Pragmatics 9(3).


Dukova-Zheleva, G. (2010). *Questions and focus in Bulgarian* (Doctoral dissertation, University of Ottawa (Canada)).


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.


