

Non-intrusive questions with matrix ‘*dali*’ in Bulgarian

VESELA SIMEONOVA¹ AND BESTE KAMALI²

¹Karl-Franzens University of Graz &

²ILLC, University of Amsterdam

ABSTRACT

This paper provides novel insights into the meaning of the particles *li* and *dali* used in matrix polar questions in Bulgarian. The primary data source is a corpus of an online chat by five family members created for the purposes of this paper. The main finding is that the two particles give rise to two different types of questions: matrix *li* questions are canonical, information seeking questions. Matrix *dali* questions, on the other hand, are non-canonical questions which we identify as non-intrusive questions in the sense of Farkas (2022) and analyze in the Table framework of Farkas & Bruce (2010).

KEYWORDS Non-intrusive questions · Non-canonical questions · Polar questions · Pragmatics

1 INTRODUCTION

Matrix polar questions (PQs) in Bulgarian are built using two alternative forms: the focus-sensitive second position clitic *li* and the sentence-initial particle *dali* (Rudin 2013). There is considerable literature on *li*, which is taken to be the canonical PQ morpheme. *Dali* has so far mostly figured in syntactic treatments of the two morphemes as the non-focus-sensitive variant merging as C (or Force in a more articulated left-periphery, e.g. Dukova-Zheleva 2010). Its meaning and whether and how it differs from the meaning of *li* have to our knowledge only been addressed in passing.

The literature mentions the meaning of *dali* in two ways: Some works implicitly assume it to be roughly semantically and pragmatically equivalent to *li* (e.g. by translating them the same way, cf. Rivero 1993, Dukova-Zheleva 2010, Romero & Han 2004). Some other researchers note in passing that *li* is the “standard” PQ form in Bulgarian while *dali* is non-standard, even “rhetorical” (Rudin et al. 1999). Rudin et al. (1999) cite a corpus study where 92.6% of neutral questions are formed with *li*. But just in what ways *dali* questions are non-standard is not explored systematically.

In this paper, we aim to clarify the meaning of *dali* and the difference between *dali* and *li*. Based on a detailed description of the usage profile of instances of *dali* occurring in a corpus we created from an online family chat, we propose a semantic generalization in terms of the absence of an expected answer. This amounts to a non-canonical, non-intrusive question in the sense of Farkas (2022), where the issue raised by the question is not expected to be resolved. We formalize our observations in the framework of Farkas & Bruce (2010) and Farkas (2022).

2 PRELIMINARIES

Both *li* and *dali* can express matrix polar questions. We refer to matrix polar questions such as those in (1) as *li* questions (1-a) and *dali* questions (1-b).¹ We approximate the meaning of the *dali* question with an added *I wonder* throughout, for reasons that will be clear in due course.

- (1) a. Prevedeni **li** sa ni pensiite?
transferred LI are our pensions
'Are our pensions transferred?'
b. **Dali** sa ni prevedeni pensiite?
DALI are our transferred pensions
'I wonder, are our pensions transferred?' [adapted from *dali*-R60]

2.1 FORMAL PROPERTIES OF *DALI* AND *LI* QUESTIONS

Syntactically, *li* is a focus-sensitive second position clitic with cognates with similar behavior in other Slavic languages (Rudin 1991, Rivero 1993, Rudin et al. 1999, Rudin 2013). Thus, (2) with *li* is a grammatical and contextually fitting question with the corresponding focus on *our pensions*. Example (3), where the associate of *li* is a topic, is contextually ill-formed. *Dali* differs starkly. It is neither a second-position element nor focus-sensitive (2). Its position is instead clause-initial (1-b) with an option to be preceded by topicalized phrases (3) (Rudin 1997, Krapova & Karastaneva 2000).

- (2) **Focus context:** Person A says that they see an income transaction in the bank statement. Person B:

[Pensiite ni]_F { **li** / #**dali** } sa prevedeni?
pensions our LI DALI are transferred
'Is it [our pensions]_F that are transferred (I wonder)?'

- (3) **Topic context:** Person A says that they are happy that they are receiving pensions as an additional income stream. Person B:

[Pensiite ni]_{TOP} { **dali** / #**li** } sa prevedeni tozi mesec?
pensions our DALI LI are transferred this month
'As for [our pensions]_{TOP}, are they transferred this month (I wonder)?'

Topicalized *li* questions are possible if the host of *li* itself is not the topic. *Dali* cannot be found in this "third" position, confirming its clause-initial (optionally post-topic) status.

- (4) [Pensiite ni]_{TOP} prevedeni { **li** / ***dali** } sa (tozi mesec)?
pensions our transferred LI DALI are this month
'As for [our pensions]_{TOP}, are they [transferred]_F this month?'

In embedded questions, *dali* introduces the embedded clause (5-b). Coupled with its (post-topical) clause-initial nature in matrix environments, this has led to analyses according to which *dali* occupies the C position in both embedded and matrix questions (Krapova & Karastaneva 2000, Dukova-Zheleva 2010, Lambova 1994). *Li* questions can also be embedded, in which case *li* may occupy any of the licit positions in relation to focus. The meanings of the two complex sentences in (5) do not significantly differ from one another.

¹Unless otherwise noted, all examples are from Bulgarian. Examples from the chat corpora are labelled with the respective Excel row (e.g. *li*-R9 means that the example is from the *li* corpus, Row 9) with their original punctuation at the end. When no reference to the corpus or outside sources is made, the example is constructed.

- (5) a. Pitax prevedeni **li** sa ni pensiite.
 asked.1SG transferred LI are our pensions
 ‘I asked whether our pensions are transferred.’
 b. Pitax **dali** sa ni prevedeni pensiite.
 asked.1SG DALI are our transferred pensions
 ‘I asked whether our pensions are transferred.’

While both *li* and *dali* are regarded as question particles, *li* has many non-interrogative uses, most comprehensively described by Rudin (1997) and attributed to *li*'s focus sensitivity, including conditionals and disjunctions, as well as phraseological expressions such as comparatives and negative adverbs (e.g. *edva li* ‘≈hardly’). *Dali*, on the other hand, is restricted to matrix and embedded interrogative environments.²

In contrast to the cross-Slavic commonality of *li*, the use of *dali* in matrix polar questions is restricted to Balkan Slavic languages. According to Rudin et al. (1999), *dali* is used rarely in neutral questions in Bulgarian whereas it has a (not fully understood) wider usage in Macedonian. Jordanoska & Meertens (2018) argue that *dali* questions are neutral questions in Macedonian based on experimental evidence (cf. also Friedman 2002, Lazarova-Nikovska 2003). Similarly, in Bosnian, Serbo-Croatian and Montenegrin, *da li* is used in neutral questions (Arsenijević 2011).

With respect to intonation, Bulgarian polar questions with both *li* and *dali* exhibit a rising contour with an early boosted peak characteristic of questions. In contrast, a typical assertion ends in a falling tune with standard local pitch accents on each prosodic word. We illustrate below with representative pitch tracks of speech produced by Author 1, who is a native speaker of Bulgarian.

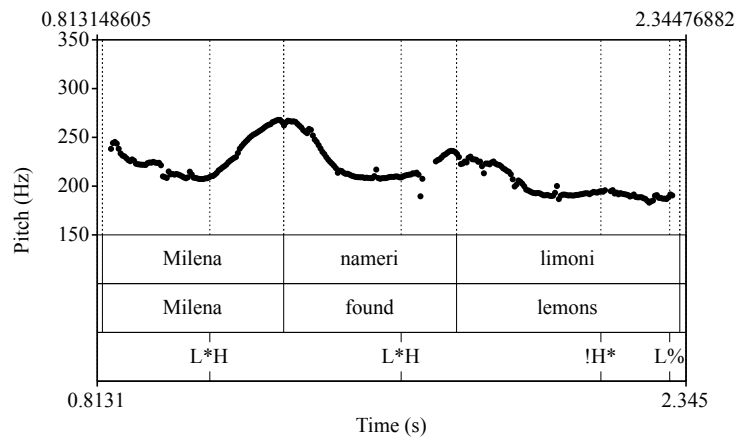


Figure 1: Assertion tune on *Milena nameri limoni* ‘Milena found lemons’

Figure 1 depicts the intonation of an assertion of the declarative *Milena nameri limoni* ‘Milena found lemons’. Our example replicates the typical assertive intonation reported by Dimitrova & Jun (2015): L*H pitch accents on each prenuclear prosodic word, and

²We take *dali* to be an atomic complementizer rather than compositionally derived from *li* and the subjunctive particle *da*, following a long tradition for Bulgarian specifically since Rudin (1993). One recent argument Krapova (2021) puts forth for this position is that *dali* and *da* can co-occur:

- (i) Ne znaeše **dali da** govori ...
 Not knew DALI DA speak
 ‘She/he didn’t know whether to speak...’ (Krapova 2021: (61))

a downstepped H* pitch accent on the nuclear word. There is a small fall after this, suggesting a L% boundary. The utterance clearly ends in a falling tune.

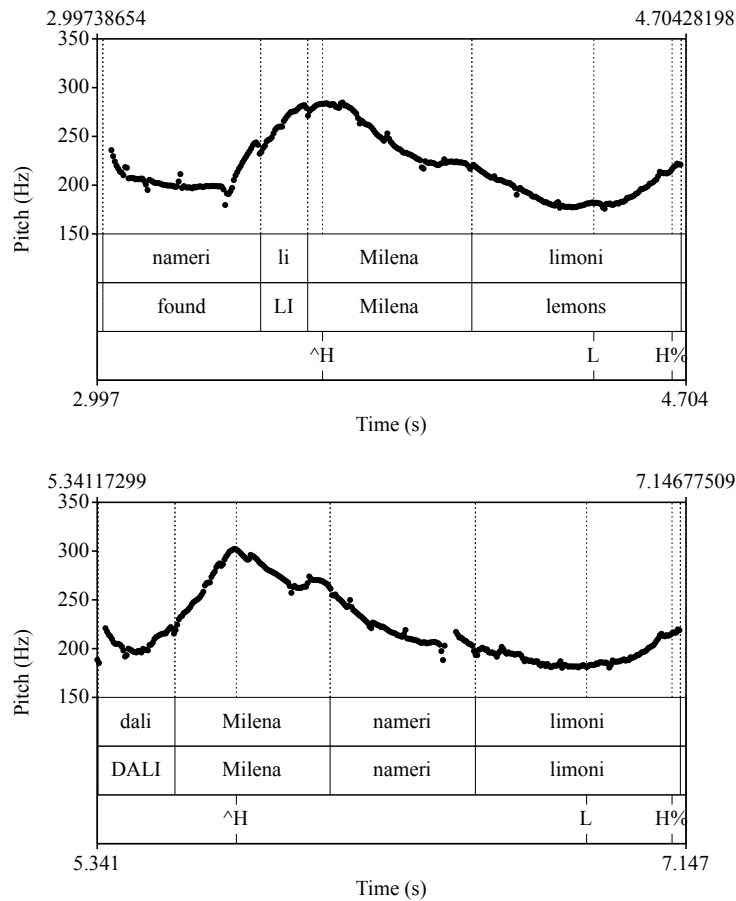


Figure 2: Question tune on *Nameri li Milena limoni* (top) and *Dali Milena nameri limoni* (bottom) ‘Did Milena find lemons’ with *li* and *dali*

In Figure 2, we see two question utterances, one with *li* (left) and one with *dali* (right). Both forms display elements of a question tune. First, they end in a clear rising boundary, a common marker of question utterances across languages, in contrast to the falling boundary of the assertion. Secondly, they feature an early ^H peak higher than the H of the first prenuclear accent of the corresponding assertion, followed by a global fall across the utterance canceling or heavily obscuring the second prenuclear accent of the assertion. Such a boosted ^H tone with surrounding tonal compression is a marker of the question tune in unrelated languages including Japanese and Turkish (Ishihara 2003, Kamali 2014). We provide conservative labels to account for the more obvious tonal pivots: upstepped ^H early in the utterance and L and H% on the final word. The intonational differences between questions with *li* and *dali* require a deeper analysis we leave to future work.

2.2 THE MEANING OF *DALI* AND *LI* QUESTIONS

Even though the literature routinely treats matrix questions with *li* and *dali* as equivalent, systematic pragmatic differences surface upon closer inspection. For example, exam

contexts, which are known to restrict admissible PQ forms (Dayal 2016), require *li*. Exam questions with *dali* are infelicitous.

- (6) (Question in history exam)
- a. Sreštat **li** se Stambolov i Bismarck?
meet.3PL LI REFL Stambolov and Bismarck
'Did Stambolov meet with Bismarck?'
 - b. #**Dali** se sreštat Stambolov i Bismarck?
DALI REFL meet.3PL Stambolov and Bismarck

In general, the use of *dali* is degraded or unacceptable in questions requesting information the addressee is assumed to be able to provide. In the contexts in (7), *Can you drive* is felicitous with *li*, but unacceptable with *dali*, because, intuitively, with *dali* the addressee is implied not to have the ability to resolve the question. Note that register is not a determining factor.

- (7) (Question in a job application form or when casually asking friends to recruit drivers for a road trip)
- a. {Umeeš / umeete} **li** da {šofiraš / šofirate}?
be.able.2SG be.able.2PL LI SUBJ drive.2SG drive.2PL
'Can you drive?'
 - b. #**Dali** {umeeš / umeete} da {šofiraš / šofirate}?
DALI be.able.2SG be.able.2PL SUBJ drive.2SG drive.2PL

Conversely, *dali* is the preferred form in a context where the addressee is presumably unable to provide an answer, such as a question about future events.

- (8) Context: a tight race where multiple contestants, Alex among them, have equal chances of winning.
- a. **Dali** Alex šte spečeli?
DALI Alex will win.3SG
'Will Alex win, I wonder?'
 - b. #Šte spečeli **li** Alex?
will win.3SG LI Alex
'Will Alex win?'

Based on data analogous to (6)-(8), Farkas (2022) has recently argued that the particle *oare* in Romanian exemplifies a type of non-canonical question that she terms *non-intrusive*. In such questions, the addressee is not expected to know or provide an informative answer. Hence, non-intrusive questions are not felicitous in information-seeking or exam contexts. In the remainder of this paper, we follow this cue and develop the proposal that *dali* questions are non-intrusive polar questions based on natural data and couch it in the framework of Farkas & Bruce (2010) and Farkas (2022).

3 THE DATA

In order to understand how *dali* is used in natural conversations, we created a corpus from a family chat and investigated actual occurrences of *dali* found in it. In this section we present examples as well as numerical and preliminary quantitative analyses that support a generalization that *dali* questions do not invite informative answers.

3.1 THE CORPUS

We created a corpus from a group chat in Viber by searching for the string [dali] and extracting all instances that were not false positive hits (e.g. the word *predali* 'conveyed'). There were 212 total true hits for *dali*, of which 53 were matrix, the focus of this paper. To

see how *li* and *dali* differ, we created a second corpus, containing a random sample of 30 matrix *li* instances.³ Due to limitations of Viber, the hits did not seem to be exhaustive.

The chat group consists of five native speakers of Bulgarian, one of whom is Author 1. They are family members living in four different countries, and the chat has been active for over ten years. Only hits dating before this research started were extracted, to avoid researcher bias.

Both corpora were annotated for response type and epistemic center (defined below) by Author 1. We believe an independent annotator was not necessary given the size and content of this study. We took measures to maintain reliability such as discussing unclear cases between authors and annotating them as unclear if certainty could not be reached. We believe that, as a trade-off, the annotation of the data by a chat participant allowed us to identify the common ground and the discourse commitments of each chat participant with more certainty. The importance of this will become clear with the examples below. For now we note that this is perhaps clearest with cases when the participants were joking, which was a significant portion of our data (about 23% of the *dali* hits). This setup provides us with a rare glimpse into the natural conversational use of *dali* and *li*.⁴

3.2 OVERVIEW OF USES OF DALI QUESTIONS

First, we present a descriptive overview of the uses of matrix *dali* in our corpus. We do not provide a similar overview for *li*, largely assumed to be unmarked, both for space reasons and because our *li* sample is representative in nature. Examples from the *li* corpus will be given in Section 3.4 to illustrate key differences between the two forms.

In our corpus, matrix *dali* questions were used predominantly in a number of non-information-seeking contexts. Their predominant function, observed in 25 occurrences, was to express that the speaker is in a state of wondering, but does not ask for an answer in the form of *yes* or *no*. These will form the core of our analysis in the following section.⁵

(9) *Wondering* (47% of examples)

- a. 'I don't remember why I didn't book the larger hotel ...'
Dali e bil mnogo sküp [...] [dali-R2]
 DALI is been very expensive
 'I wonder, was it very expensive?'
- b. (Uttered in October 2020, when the stock market was turbulent due to Covid.) 'What might have happened with my investments ...'
Dali sa padnali? [dali-R96]
 DALI are fallen
 'I wonder, have they fallen?'
- c. 'My vaccination was moved from 31.01 to 18.02 ...'
Dali šte ima njakaküv efekt? [dali-R79]
 DALI will have.3SG some effect
 'I wonder, will it have any effect?'
- d. 'The quarantine in France is until 24.07...'
Dali šte može da pütuvate togava? [dali-R130]
 DALI will can SUBJ travel.2PL then
 'I wonder, will it be possible for you to travel then?'

³*Li* instances proved difficult to count due to many unrelated strings of [li] in the data and the limitations of the chat environment. We presume that they are overall more numerous than the instances of *dali*.

⁴Due to the nature of the data – a private conversation among identifiable participants – we were unable to make the corpus public. We did, however, obtain permissions from all participants to include the unaltered examples in this paper, for which we are grateful. Author 1 remains available by personal communication to address questions regarding the corpus.

⁵Note that the first example is shortened here, its full version is in (16).

- e. (The speaker is about to travel and is debating which jacket to wear.)
Dali njama da mi e studeno s červenoto?
 DALI not SUBJ me.DAT is cold with red.DEF
 'I wonder, won't I be cold with the red one?' [dali-R165]
- f. (Uttered by the only person in the group who plays the lottery, so no one else is likely to know the answer.)
 Tototo **dali** raboti, če imam da vzemam pečalba?
 lottery DALI works that have.1SG SUBJ collect.1SG winnings
 'I wonder, is lottery shop open, because I have winnings to collect?'
 [dali-R150]

The next most popular category was humorous uses, which comprised 12 occurrences. Notice that some of these uses are rhetorical in the sense that the answer is taken to be obvious to all (see Biezma & Rawlins 2017:and references therein).

(10) *Humor* (22.6% of examples)

- a. 'Hey folks, ever since I bought a new shampoo and a conditioner (online), I started getting ads for hair products ...'
Dali e slučajnost?
 DALI is coincidence
 'I wonder, is this a coincidence?' [dali-R116]
- b. (The speaker is over 60 years old.) 'At the [someone else's birthday] party, I wished to be 60 again ...'
Dali šte mi se sbūдне na men?
 DALI will me.DAT REFL happen to me
 'I wonder, will my wish come true?' [dali-R163]
- c. (Referring to a photo of two star ratings on the door of a communist-style restaurant, where the rating is out of five stars – unlike Michelin where it is out of three.)
 Dvete zvezdi **dali** sa Michelin? (wink)
 two.DEF stars DALI are Michelin
 'I wonder, are the two stars Michelin?' [dali-R174]

There were also 5 uses expressing wishes or hopes, and 3 uses expressing polite requests or signalling increased politeness more broadly.

(11) *Wishes, hopes* (9.4% of examples)

- a. (Referring to a picture of something sci-fi looking. The addressee occasionally writes sci-fi short stories.)
 Tate, **dali** tova može da ti posluži za vdūhnovenie
 dad.VOC DALI this can SUBJ you.DAT serve.3SG for inspiration
 za nov razkaz?...
 for new short.story
 'Dad, I wonder, could this serve you as inspiration for a new short story?' [dali-R11]
- b. (Uttered in March 2020, at the outset of the Covid-19 pandemic, and the defense was scheduled for late March)
 Vessie, a tvojata zaštita **dali** šte se sūstoi po plan?
 name and your defense DALI will REFL happen on plan
 'Vessie, and your defense, I wonder, will it happen according to plan?'
 [dali-R154]

- c. (The addressee had ordered a new bed for her daughter but the delivery was delayed multiple times.)
Dali šte dojde novoto leglo za ČRD
 DALI will arrive new bed for birthday
 ‘I wonder, will the new bed make it before [your daughter’s] birthday?’
 [dali-R98]
- (12) *Politeness* (5.6% of examples)
- a. (The addressee is in Bulgaria, the speaker is not) ‘I cannot find my sandals...’
 Mamče, **dali** sa ostanali v Bg.
 mom DALI are remained in Bulgaria
 ‘Mom, I wonder, did I leave them in Bulgaria?’ (≈ ‘Can you check?’)
 [dali-R212]
- b. Hora **dali** može da vi pomolja za usluga.
 folks DALI can SUBJ you.ACC ask.1SG for favor
 ‘Folks, I wonder, could I ask you a favor?’ (next sentence: ‘A friend of mine is looking for data in Bulgarian.’)
 [dali-R211]
- c. (An acquaintance is travelling from the speaker’s country to the addressee’s. The acquaintance can be asked to bring along some small things, but cheese and phyllo are probably too bulky, and the addressee will probably not approve of asking the acquaintance to bring them along.)
 I, **dali** da te popitam, ne smeja. Da pratja
 name DALI SUBJ you.ACC ask.1SG not dare.1SG SUBJ send.1SG
 malko kaškaval... kori za banica nešto
 little cheese pastry.phyllo something
 ‘I, I wonder, should I ask you, [but] I dare not. Could I send you a bit of cheese... or phyllo, or something?’
 [dali-R22]

In summary, most *dali* instances we find in our corpus can be collapsed under the usage categories of wondering, humor, wishes, or politeness. Even intuitively, these are non-information-seeking contexts where the speaker does not expect a *yes* or *no* answer from the addressee(s), and in the next two sections, we substantiate the claim that *dali* questions are indeed not information-seeking by showing that they neither require informative answers from the other interlocutors, nor expect epistemic authority from them.

3.3 RESPONSE BEHAVIOR

Questions that give rise to a wondering flavor are often described as obviating the expectation that the addressee can answer these questions (Truckenbrodt 2006, Eckardt 2020, Farkas 2022). In standard information-seeking questions, on the other hand, the speaker wants to learn the answer and either knows that the addressee can provide the answer, or at least assumes that there is a possibility that they do. Such an asymmetry may be empirically verified by investigating the number and character of responses to each type of question in natural discourse.

To determine the response behavior of *dali* questions and *li* questions, we counted instances of informative answer, no answer, non-definitive answer (such as “maybe”, “I don’t know”, “hopefully”, or guesses) and cases where no answer was needed (such as jokes) in our corpora. We expected *li* questions to lead to informative answers more frequently than *dali* questions.

In Table 1, we present our findings in terms of raw count, percentage, and unpaired t-test result. The differences between the values listed above the line are statistically significant, and below the line they are not.

As expected, *dali* questions received an informative answer only 13.2% of the time. 58.5% of the reactions to *dali* were no answer at all, with a further 28.3% comprising

Response type	<i>li</i>	<i>dali</i>	<i>p</i> -value
informative answer	20/30 (66.6%)	7/53 (13.2%)	<0.0001
not answered	5/30 (16.6%)	31/53 (58.5%)	0.0001
not definitive	N/A	15/53 (28.3%)	0.0010
not needed	7/30 (23.3%)	12/53 (22.6%)	0.9434

Table 1: Whether the question received an informative answer and what kind.

non-definite answers. In contrast, *li* questions were answered informatively 66.6% of the time. The absence of an answer was the exception, and our sample did not include any non-definite answers. The category of questions where no answer was needed, in which the two forms did not differ from one another, are mostly comprised of humorous and sarcastic uses. We believe this finding is not informative in terms of answer behavior, but rather in terms of the fact that the humorous function can be served both by *li* and *dali*.

We can thus verify empirically that *dali* questions are routinely left unanswered, while *li* questions are mostly answered informatively in actual discourse. This response behavior confirms *li* questions as canonical information-seeking questions. Conversely, the typical response behavior of *dali* questions, which were mostly left unanswered, supports the view that these questions do not expect informative answers.

3.4 EPISTEMIC CENTER

The previous section showed, based on the observed pattern of actual responses, that while *li* questions are most often answered, *dali* questions are most often unanswered. Here we tackle the more elusive intuition of whether the speaker in fact expects an answer. We reason that the speaker does not expect an answer if they already know that no addressee in the conversation has the knowledge necessary to answer the question. To track this kind of information we must identify the *epistemic center*: the interlocutor who has – or is reasonably likely to have – the information to answer the question.

To determine how *li* and *dali* questions in our corpora may differ in terms of epistemic center, we analyzed each instance with respect to who has the information to answer the question, and subjected the resulting counts to a t-test. The annotation labels we used are defined as follows:

- (13) Epistemic Center labels used in annotation
- a. **addressee**: when the addressee is assumed to know the answer to the question (this is most obviously the case if the question is about the addressee)
 - b. **speaker**: when the speaker would know the answer best
 - c. **nobody**: when no conversation participant can possibly know the answer (e.g. in questions about what will happen in the future)
 - d. **everybody**: when the answer is already known by all conversational participants (e.g. in the case of jokes)
 - e. **unclear**: when it was not clear (based on the common ground) who the epistemic center is

Table 2 summarizes the proportions of each type of epistemic center as well as the results of an unpaired t-test, again plotting statistically significant differences between the two forms above the line.

The results show that a remarkable majority of the data points cluster under two of the epistemic center parameters. The first is when the addressee is the epistemic center: while 73.3% of *li* questions fall in this category, *dali* questions are only rarely (3/53 times) used in this situation. (14) provides examples from the *li* corpus in this most typical use, where the addressee is clearly the epistemic authority on whether they got some sleep or whether the flower in their room smells nice.

Epistemic center	<i>li</i>	<i>dali</i>	<i>p</i> -values
addressee	22/30 (73.3%)	3/53 (5.6%)	<0.0001
nobody present	0/30 (0%)	20/53 (37.7%)	<0.0001
speaker	2/30 (6.6%)	6/53 (11.3%)	0.4960
everybody present	4/30 (13%)	9/53 (16.9%)	0.6651
unclear	2/30 (6.6%)	7/53 (13.2%)	

Table 2: Who was the epistemic center? Above the line=statistically significant; below=not.

(14) *Li* questions with addressee as the epistemic center

- a. Pospa **li** v samoleta?
slept.a.bit LI in plane.DEF
'Did you get some sleep on the plane?' [li-R16]
- b. (Replying to a photo of the addressee holding an ice-cream cone.)
Šokoladovo **li** e?
chocolate LI is
'Is it chocolate [flavor]?' [li-R12]
- c. (Replying to a photo of the addressee's hyacinth plant starting to bloom.)
Uhae **li**?
smell.nice li
'Does it smell nice [already]?' [li-R30]
- d. (Commenting on a photo of a group of people, including the addressee.)
Snimkata ot sega **li** e?
photo.DEF from now LI is
'Is the photo [very] recent?' [li-R4]

Dali questions, on the other hand, are most frequently used when neither the speaker nor the addressee(s) know the answer. Many of the *dali* examples with a wondering flavor given in (9) illustrate this use, for example, making reference to unknowable future events such as the weather or whether the government would allow travelling, etc. Notably, there were no instances of *li* in this category, suggesting that *li* requires some epistemic authority of the addressee, whereas *dali* does not.

If we briefly evaluate the statistically insignificant results in Table 2, we again note the shared function of humor, which accounts for the majority of cases with epistemic center 'everyone'. Examples of this sort with *dali* are provided in (10). (15) is an example with *li*. Here, both the speaker and the direct addressee know very well that what is meant by the addressee is not fashion reviews.

- (15) (In response to the Addressee's statement 'I have many review deadlines'. Both primary interlocutors are academics.)
Modni **li**?
fashion li
'Fashion reviews?' [li-R8]

Secondly, we see that both *li* and *dali* may be used when the epistemic center is the speaker. We suggest that *li* questions with the speaker as the epistemic center exemplify true self-addressed questions. Here, we assume along with Farkas (2022), that such self-addressed questions require the speaker to effectively simultaneously be the addressee. We leave the contrasts between *li* and *dali* used in these specific shared functions to future work.

To summarize, the results show that matrix *li* questions are predominantly used when the addressee is assumed or known to hold information that answers the question. *Dali*

questions, on the other hand, are used when nobody, or nobody other than the speaker, is assumed to have information to answer the question at hand. Hence, the differential use of *li* and *dali* based on epistemic center once again supports the central claim of this paper, that matrix *dali* and *li* questions are not equivalent: *li* questions are canonical, information-seeking questions, while *dali* questions are used when the speaker expects no informative answer to their question. This integral part of the pragmatics of *dali* is formalized in the next section.

4 THE PROPOSAL

In this section, we formalize the observation that *dali* questions do not seek an informative answer by analyzing them within the Table framework (Farkas & Bruce 2010) as non-intrusive questions in the sense of Farkas (2022).

4.1 A FEW OBSERVATIONS

Our proposal that *dali* questions do not seek an informative answer predicts differences in their felicity with first person subjects (or more broadly, when the speaker is involved in the action). We saw a *dali* question with a first person subject in example (9-a) repeated below in (16). In this example, the speaker states that he does not remember why he booked one hotel for his vacation and not another, and proceeds to contemplate some likely answers (e.g. the price difference). Since this was his personal vacation that he organized himself, there is no expectation that anybody else knows the answer (though it is not impossible, e.g. if he had shared his thoughts when he was booking).

- (16) 'I don't remember why I didn't book the larger hotel...'
- Dali** e bil mnogo sküp (ili ne e imalo mesta, ili ne možex da
DALI is been very expensive or not is had spots or not could SUBJ
gi kupja online...)?
them buy online
'I wonder, was it very expensive (or there weren't any spots left, or I couldn't
book them online...)?' [dali-R2]

If this question is formed with *li* instead, as in (17), the interpretation is that the speaker assumes that the addressees are in a better position to know the reason for his action. This would only be the case if he once told someone in the group why he didn't book the other hotel, or assumes so. Our approach accounts for this intuition.

- (17) Mnogo sküp **li** e bil?
very expensive LI is been
'Was it very expensive?'

Similarly, we predict that expecting an informative response from the addressee is not part of the not-at-issue (NAI) meaning of *dali*, but it is for *li*. This can be verified by using the *actually* test introduced by Francis (2021); *actually* challenges NAI content, in particular, discourse expectations. An addressee cannot reply to a canonical information-seeking question with '*actually, I know the answer*' because the expectation that the addressee knows the answer is part of the NAI content of information-seeking questions.

- (18) A: Did the Leafs lose the game last night?
B: Actually, I have no idea.
B' #Actually, I know whether they did. English; after Francis (2021)

This test applies to *li* and *dali* questions as follows: we predict that if *li* questions are indeed information-seeking, the felicity of the use of *actually* would mirror that of the information-seeking (18). In contrast, with *dali*, the felicity values will be reversed

because with *dali* the default expectation of the speaker (albeit cancellable) is that the addressee does not have an answer. This is exactly what we observe.

- (19) A: Snimkata ot sega **li** e?
photo from now LI is
'Is the photo recent?'
B: Vsüştnost, ne znam. 'Actually, I don't know.'
B' #Vsüştnost, znam dali e ot sega. 'Actually, I know if it is.'
- (20) A: **Dali** sa ni prevedeni pensiite?
DALI are our transferred pensions
'I wonder, are our pensions transferred?'
B: #Vsüştnost, ne znam. 'Actually, I don't know.'
B': Vsüştnost, znam dali sa prevedeni. 'Actually, I know if they are transferred.'

In the rest of this section, we propose that while matrix *li* questions are canonical, matrix *dali* questions are non-canonical, namely non-intrusive in the sense of Farkas (2022). The next section presents the framework and the formal proposal.

4.2 THE TABLE FRAMEWORK AND *LI* QUESTIONS

According to Farkas & Bruce (2010), utterances in conversation have two main components: For one, they make private commitments public – for example, when A utters p ='it is raining' A publicly commits to p ; this is called a Discourse Commitment of A (DC_A). The Common Ground is the set of propositions that have been agreed upon by all participants, together with the propositions that represent the shared background knowledge of the discourse participants.

Secondly, utterances steer the conversation: an utterance is a proposal to update the common ground (e.g. the prototypical assertion of p proposes to update the CG with a joint commitment to p); this is called the Projected Set (PS).

Conversational acts are formally represented in a workspace called the Table. Table 3 shows the context structure after the utterance of a proposition p .

$DC_{Speaker}$	Table	$DC_{Addressee}$
$info(I)=p$	$\{p\}$	
	$PS = \{DC_{Ad} \cup \{p\}\}$	

Table 3: Context structure after the Speaker has uttered an assertion, e.g. 'It is raining,' in the framework of Farkas & Bruce (2010)

The denotation of utterances is called an Issue. Issues have informative content, marked with $info(I)$ following Farkas (2022). The issue is the union of the propositions in I . In a (non-modal) canonical assertion like *It is raining*, $info(I)=p$ (where $p=it\ is\ raining$), i.e. the informative content of the issue is a singleton proposition (partitioning the logical space W and denoting only the set of worlds where it is raining). By uttering p , the Speaker proposes that p be added to the Discourse Commitments of the Addressee.

Following the consensus in the literature as well as our findings reported in the previous section, we analyze Bulgarian *li* questions as canonical information-seeking questions. Attached to the verb as in (21), *li* expresses broad focus. Such a canonical information-seeking question is illustrated in Table 4 in the formal Table representation.

- (21) *Li* questions as canonical polar questions
a. Prevedeni **li** sa ni pensiite?
transferred LI are our pensions
'Are our pensions transferred?'

b. $Issue = \{p, \neg p\}$

$DC_{Speaker}$	Table	$DC_{Addressee}$
$info(I)=W$	$\{p, \neg p\}$	
	$PS = \{DC_{Ad} \cup \{p\}, DC_{Ad} \cup \{\neg p\}\}$	

Table 4: Context structure after the Speaker has uttered a canonical polar question (21-a), in the framework of Farkas & Bruce (2010)

In a canonical polar question, the informative content of the issue contains two alternatives, p and $\neg p$. Because questions do not commit their author to either of the alternatives generated by the radical, the DC of the Speaker is trivial, i.e. W is not partitioned. Canonical polar questions project an inquisitive context with respect to their sentence radical: the PS contains both future common grounds to which p is added and future common grounds to which $\neg p$ is added. In a future discourse state, the addressee's response will determine which will be added to the CG and the issue will be resolved.

4.3 DALI QUESTIONS AS NON-INTRUSIVE QUESTIONS

Farkas (2022) notes that canonical questions, such as English polar interrogatives and Bulgarian *li* questions, have the following default assumptions.

- (22) Default assumptions accompanying question acts (Farkas 2022: 297)
- a. *Speaker ignorance*: The speaker's epistemic state is neutral relative to the possible resolutions of the issue she raises.
 - b. *Addressee competence*: The speaker assumes that the addressee knows the information that settles the issue she raises.
 - c. *Addressee compliance*: The speaker assumes that the addressee will provide this information in the immediate future of the conversation as a result of the speaker's speech act.
 - d. *Issue resolution goal*: It is assumed that the main aim the speaker pursues when raising an issue is to have it resolved in the immediate future of the conversation.

Non-canonical questions are questions where one or more of these default assumptions are suspended. In English tag questions, for example, the speaker ignorance assumption is suspended.

The Romanian particle *oare* provides another case of non-canonical question. As noted in Section 2.2, *dali* has numerous parallelisms with *oare*, including a wondering flavor, infelicity in strictly information-seeking contexts, and preferred usage in contexts where the addressee may not know the answer.

- (23) (There is a knock on the door in the middle of the night. Maria says to Paul:)
- ?(Oare) cine e la ora asta?
 OARE who is at hour this
 'Who could it be at this hour?' Romanian (Farkas 2022: (20))

Farkas (2022) analyzes these cases as suspending the assumption of addressee compliance (entailing addressee competence). For example, in (23), the addressee is assumed not to be able to provide an answer, or as a device that "blunts the 'putting the addressee on the spot' effect of questions" (Farkas 2022: 312).

Following the characterization of non-canonical moves by Farkas & Roelofsen (2017), Farkas (2022) postulates a Special Discourse Effect to account for the meaning of *oare*

questions. The effect consists in adding $\{\text{info}(I)\}$ to the projected set. The Addressee may choose to “resolve” the issue by adopting the unresolved issue itself, i.e. by joining the wondering state. Such non-canonical questions are termed *non-intrusive questions* by Farkas and defined in (24) and represented in Table 5.

- (24) *Non-intrusive questions* (Farkas 2022: (38))
 A question is non-intrusive iff the PS of its output context state includes $DC_X \cup \{\text{info}(I)\}$, where I is the issue placed on the Table by the question.

We propose that *dali* questions are non-intrusive polar questions in the sense of Farkas (2022).

- (25) *Dali* questions as non-canonical polar questions with a special discourse effect
- a. **Dali** sa ni prevedeni pensiite?
 DALI are our transferred pensions
 ‘I wonder, are our pensions transferred?’
 - b. *Issue* = $\{p, \neg p\}$
 - c. Special discourse effect: $PS \cup \{DC_{Addressee} \cup \{\text{info}(I)\}\}$

$DC_{Speaker}$	Table	$DC_{Addressee}$
$\text{info}(I)=W$	$\{p, \neg p\}$	
	$PS = \{DC_{Ad} \cup \{p\}, DC_{Ad} \cup \{\neg p\}, DC_{Ad} \cup \{\text{info}(I)\}\}$	

Table 5: Context structure after the Speaker has uttered a non-intrusive polar question in the framework of Farkas (2022)

The Special Discourse Effect in (25-c) consists in not only enriching the Projected Set with either of the two polar alternatives, but also the possibility that the Projected Set remains unchanged (not enriched), by being updated with only *Info(I)*. This addition makes it possible for the Addressee not to go beyond the Speaker’s current commitments. The possibility for the lack of a content update on the CG makes *dali* questions non-canonical: in uttering a *dali* question, the Speaker is signalling that the projected CG need not be enriched with the commitments *Our pensions are transferred* and *Our pensions are not transferred*.

In summary, we analyze the contrast between *li* and *dali* matrix polar questions as canonical and non-intrusive, non-canonical questions, respectively, capturing the novel observations about their distribution differences reported in this paper.

5 CONCLUDING REMARKS

In this paper, we have provided an empirical description and novel analysis of Bulgarian matrix *dali* questions. The naturally occurring data reported here demonstrate that unlike *li* questions, *dali* questions do not expect an informative answer from the addressee. This matches the notion of non-intrusive question, where the assumption of addressee compliance in providing an answer is suspended, in the framework of Farkas & Bruce (2010) and Farkas (2022).

Our findings show a broader crosslinguistic applicability of the category of non-intrusive questions and its analysis *à la* Farkas (2022). Our treatment goes even further in the empirical realm, using natural data and establishing quantifiable markers of non-intrusiveness such as response behavior and epistemic center. Furthermore, we have uncovered data with various distinct uses, which will be important to better understand this category.

Although we have adopted Farkas' analysis in its entirety, there are questions to investigate in more detail that fall outside of this paper's scope. One important difference between Bulgarian *dali* and Romanian *oare* is that *dali* is the canonical embedded polar question complementizer and does not give rise to special effects in embedded environments (recall (5-b)). For this reason, unlike Romanian non-intrusive questions, Bulgarian non-intrusive questions may be said to exhibit *insubordination*, i.e. a subordinate form acting as a matrix form (Evans 2007). Insubordinate questions are attested in German, known as verb-final questions (Oppenrieder 1989, Truckenbrodt 2006).

- (26) **Ob** Katelbach wohl kommt?
 whether Katelbach indeed comes
 'Will Katelbach come, I wonder?' German (Oppenrieder 1989: (57))

Eckardt (2020) analyzes German verb-final questions such as (26) as conjectural questions in the sense of Littell et al. (2010), i.e. involving conjectural evidentials. While conjectural questions are very close to non-intrusive questions in meaning, Farkas (2022) argues that they are not identical. *Dali* provides independent support for this claim, since *dali* is not an evidential marker. Furthermore, the formal parallelism between *dali* questions and German verb-final questions opens up the possibility that insubordination itself may feed a subset of non-intrusive/conjectural characteristics in the absence of an evidential.

Another important difference between *oare* and *dali* is that *oare* is used in both polar and content non-intrusive questions, while *dali* is banned from *wh*-questions. Non-intrusive *wh*-questions exist in Bulgarian (and other South Slavic languages), but they are formed with *li* (see Simeonova to appear) and *dali* is banned from them, which is surprising in light of the findings in this paper. This poses a compositional challenge that we leave to future work.

ACKNOWLEDGMENTS

We thank the audience at FASL 31 for their questions and especially the anonymous reviewers for their feedback, which has greatly improved this paper. We also thank the editors of this issue as well as of JSL for their support through the publishing process. We are grateful to Patrick Grosz, Edgar Onea, and Radek Šimik for helpful discussions. We thankfully acknowledge the financial support of FWF grant I 4858-G (Vesela Simeonova; PI Edgar Onea) and EU grant MSCA 101067203-EPOQ (Beste Kamali).

CONTACT

VESELA SIMEONOVA — vesela.simeonova@uni-graz.at
 BESTE KAMALI — b.c.kamali@uva.nl

ABBREVIATIONS

ACC	accusative	PL	plural
AUX	auxilliary	PS	projected set
CG	common ground	REFL	reflexive
DAT	dative	SG	singular
DC	discourse commitment	SUBJ	subjunctive
DEF	definite	TOP	topic
F	focus	VOC	vocative
NAI	not-at-issue		

REFERENCES

- Arsenijević, Boban. 2011. Serbo-Croatian coordinative conjunctions at the syntax-semantics interface. *The linguistic review* 28(2). 175–206.
- Biezma, Maria & Kyle Rawlins. 2017. Rhetorical questions: Severing asking from questioning. In Dan Burgdorf, Jacob Collard, Sireemas Maspong & Brynhildur Stefánsdóttir (eds.), *Proceedings of SALT 27*, 302–322.
- Dayal, Veneeta. 2016. *Questions*. New York: Oxford University Press.
- Dimitrova, Snezhina & Sun-Ah Jun. 2015. Pitch accent variability in focus production and perception in Bulgarian declaratives. In *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow: University of Glasgow.
- Dukova-Zheleva, Galina. 2010. *Questions and focus in Bulgarian*: University of Ottawa dissertation.
- Eckardt, Regine. 2020. Conjectural questions: The case of German verb-final *wohl* questions. *Semantics and pragmatics* 13(9). 1–47.
- Evans, Nicholas. 2007. Insubordination and its uses. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. New York: Oxford University Press.
- Farkas, Donka. 2022. Non-intrusive questions as a special type of non-canonical questions. *Journal of semantics* 39(2). 295–337.
- Farkas, Donka & Kim Bruce. 2010. On reacting to assertions and polar questions. *Journal of semantics* 27(1). 81–118.
- Farkas, Donka & Floris Roelofsen. 2017. Division of labor in the interpretation of declaratives and interrogatives. *Journal of semantics* 34(2). 237–289.
- Francis, Naomi. 2021. *Actually* and the marking of (un)controversiality. Presented at the LingORG colloquium at the University of Oslo, March 15, 2021.
- Friedman, Victor. 2002. *Macedonian*. Munich: Lincom Europa.
- Ishihara, Shinichiro. 2003. *Intonation and interface conditions*: Massachusetts Institute of Technology dissertation.
- Jordanoska, Izabela & Erlinde Meertens. 2018. The pragmatic effects of Macedonian *li*: an empirical study. In Andreas Blümel, Jovana Gajić, Ljudmila Geist, Uwe Junghanns & Hagen Pitsch (eds.), *Advances in formal Slavic linguistics*, Berlin: Language Science Press.
- Kamali, Beste. 2014. Beyond morphosyntax: Interrogative intonation and its role in Turkish. *Turkic languages* 18(1/2). 189–206.
- Krapova, Iliyana. 2021. Complementizers and particles inside and outside of the left periphery: The case of Bulgarian revisited. In Björn Wiemer & Barbara Sonnenhauser (eds.), *Clausal complementation in South Slavic*, 211–269. Berlin: De Gruyter.
- Krapova, Iliyana & Tsena Karastaneva. 2000. Complementizer positions in Bulgarian. In Anisava Miltenova (ed.), *Bulgarian studies at the dawn of the 21st century*, 93–104. Sofia: Gutenberg.
- Lambova, Mariana. 1994. *On information structure and clausal architecture: Evidence from Bulgarian*: University of Connecticut dissertation.

- Lazarova-Nikovska, Ana. 2003. On interrogative sentences in Macedonian: A generative perspective. *RCEAL working papers in English and applied linguistics* 9. 129–159.
- Littell, Patrick, Lisa Matthewson & Tyler Peterson. 2010. On the semantics of conjectural questions. In *Evidence from evidentials*. University of British Columbia working papers 28, 89–104. Vancouver: UBC.
- Oppenrieder, Wilhelm. 1989. Selbständige Verb-letzt-Sätze: Ihr Platz im Satzmodussystem und ihre intonatorische Kennzeichnung. In Hans Altmann, Anton Batliner & Wilhelm Oppenrieder (eds.), *Zur Intonation von Modus und Fokus im Deutschen*, 163–244. Tübingen: Niemeyer.
- Rivero, Maria-Luisa. 1993. Bulgarian and Serbo-Croatian yes-no questions: V⁰-raising to *-li* versus *-li* hopping. *Linguistic inquiry* 24(3). 567–575.
- Romero, Maribel & Chung-Hye Han. 2004. On negative yes/no questions. *Linguistics and philosophy* 27(5). 609–658.
- Rudin, Catherine. 1991. Topic and focus in Bulgarian. *Acta linguistica hungarica* 40(3/4). 429–447.
- Rudin, Catherine. 1993. On focus position and focus marking in Bulgarian questions. In *Papers from the 4th annual meeting of the Formal Linguistics Society of Midamerica*, 252–265.
- Rudin, Catherine. 1997. Kakvo li e li: Interrogation and focusing in Bulgarian. *Balkanistika* 10. 335–46.
- Rudin, Catherine. 2013. *Aspects of Bulgarian syntax: Complementizers and WH constructions*. Bloomington: Slavica.
- Rudin, Catherine, Christina Kramer, Loren Billings & Matthew Baerman. 1999. Macedonian and Bulgarian *li* questions: Beyond syntax. *Natural language & linguistic theory* 17(3). 541–586.
- Simeonova, Vesela. to appear. Mirative and deliberative non-canonical content questions in South Slavic. *Journal of Slavic linguistics*.
- Truckenbrodt, Hubert. 2006. On the semantic motivation of syntactic verb movement to C in German. *Theoretical linguistics* 32(3).