Bulgarian *da* as a non-indicative placeholder

Hagen Pitsch (Göttingen)

1. Introduction

This paper is concerned with the interpretation of Bulgarian *da*-constructions and the semantics of *da*. *Da*-constructions are (at least) made up of the uninflected element *da* and a verb phrase (VP) in an indicative (IND) tense form; cf. (1):

\[ \text{[XP} \text{da}_X \ldots \text{[VP}_{\text{IND}} (\text{VP})]\]

While the syntax of Bulgarian *da*-constructions seems thus rather simple (though more has to be said about XP), their meaning is less obvious, which is the reason to address it once again. There is a long-standing debate in the literature on whether or not *da*-constructions form the Bulgarian subjunctive mood. Authors denying this view either (i) argue that Bulgarian lacks such a category altogether (e.g., Genadieva-Mutačieva 1970, 1976), or (ii) suggest that *da*-constructions may function not only as subjunctives, but also as infinitivals (e.g., Popov 1968; Lempp 1981; Krapova 1998, 2001), or (iii) propose a varying number of homonymous *da*-constructions with diverse functions and meanings (e.g., Petkova Schick 1977; Maslov 1981; Tilkov et al. 1983; Ivanova 2014).

The goal of this paper is to account for the interpretation of Bulgarian *da*-constructions as well as to support the view that Bulgarian has not only subjunctive, but also infinitival CPs. I will argue that *da* marks the absence of the indicative with its characteristic supposition of *subject certainty* (Siegel 2009). Furthermore, I will assume that Bulgarian *da*-constructions form a semantic default used when more specific moods are blocked as they

---

1 Thanks are due to Andreas Blümel, Eva Csipak, Uwe Junghanns, Frank Sode, Viola Schmitt and Hedde Zeijlstra – for listening to my ideas and for valuable advice – and to the audience of FDSL 11 for a helpful discussion. I am also indebted to Krăstina Arbova-Georgieva for important suggestions and help with the data.

2 Indicative tenses report situations as (to be) realized. Among them, the present, present perfect, past perfect and (seldom) the imperfect occur in *da*-constructions, while the aorist, the future and future perfect do not (cf. Krapova 1998: 81). The past future and past future perfect are non-indicative tenses and resemble *da*-constructions in reporting situations as hypothetical/unrealized. But other than *da*-constructions, which often show a ‘possible future’ interpretation (Krapova 1998: 81–2), they relate to (a relative) past.

3 Among the authors holding this opinion are, e.g., Weigand (1907), Seliščev (1952), Maslov (1962), Bernštejn (1961), Kramer (1992), Smirnova (2010, 2012).

4 This paper thus adds to the discussion led in, a.o., Maslov (1962), Genadieva-Mutačieva (1970), Petkova Schmidt (1977), Siegel (2009), Smirnova (2012).
would cause a semantic failure in the given context (cf. Schlenker 2005). Importantly, this view still allows the distinction of subjunctive-like and infinitive-like da-constructions (cf. Popov 1968; Krapova 1997, 1998, 2001), since this distinction is not based on the (semantically vacuous) particle da, but rather on the properties of the Tense complex of the relevant da-construction. Consequently, da can unambiguously be identified with the category of mood and simultaneously be dissociated from the notion of finiteness.

The paper is structured as follows: Section 2.1 recapitulates Siegel’s (2009) analysis on the semantics of Balkan subjunctives. In Section 2.2, I argue for a distinction of finite vs. non-finite da-constructions in Bulgarian (cf. Krapova 1998). In Section 2.3, the priorly introduced ingredients are combined into the proposal proper. Section 2.4 serves to illustrate my assumptions concentrating on da-complements. Section 3 will summarize the paper.

2. The proposal
2.1 Semantic default and subject certainty

In this section, I recapitulate Siegel’s (2009) proposal on the semantics of mood in Balkan languages, among them Bulgarian, which will be the first ingredient of my proposal. Siegel argues that, ‘while indicative morphology is specified as being [+realis], subjunctive morphology is underspecified for semantic content, and appears as a default when other, more specified, “moods” cannot appear.’ (Siegel 2009: 1879) Assuming Bulgarian da-constructions to be subjunctives, this amounts to the claim that they do not contribute to sentence semantics. It follows that they represent the neutral choice within the Bulgarian mood system, a choice employed whenever the remaining moods are too specific.4 As Siegel (2009) points out, ‘treating subjunctive as a default in this way is appealing because subjunctive morphology appears in such a wide range of environments, environments for which it has proved to be very difficult to provide any unified analysis.’ (Siegel 2009: 1879)

The second important insight of Siegel’s analysis is that, ‘in Balkan, indicative is correlated with a higher degree of certainty on the part of the subject than is subjunctive.’ (Siegel 2009: 1879) Note that this does not mean that the subjunctive is associated with a ‘weaker’ indicative semantics, since the mentioned correlation is a relative one and based on

---


5 Smirnova (2012) is an attempt for such an analysis. Smirnova assumes the Bulgarian subjunctive to presuppose weaker epistemic commitment than the indicative. More precisely, it is claimed to presuppose that the domain with respect to which the relevant proposition is evaluated is non-homogenous, while the indicative presupposes a homogenous domain. Importantly, Smirnova assumes that both the indicative and the subjunctive are meaningful mood forms, while Siegel’s (2009) account allows the latter to be a semantically vacuous default.
the fact that the subjunctive is in opposition to the indicative. In fact, the subjunctive is void of semantic content. From the discussed assumptions, it is possible to preliminarily outline the mood system of modern Bulgarian (imperative omitted):

<table>
<thead>
<tr>
<th></th>
<th>indicative</th>
<th>subject certainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>∅</td>
<td></td>
<td></td>
</tr>
<tr>
<td>da</td>
<td>non-indicative</td>
<td>——</td>
</tr>
</tbody>
</table>

Table 1: Mood in Bulgarian (I).

2.2 Finite and non-finite da-constructions

The second ingredient of my proposal is the claim that da-constructions happen to occur in two varieties, namely finite and non-finite da-constructions. In her paper on subjunctive complements in Bulgarian, Krapova (1998) (see also Krapova 1997, 2001) makes the observation that Bulgarian has two types of S[subjunctive]s (i.e. da-constructions) which she calls ‘Type I Ss’ and ‘Type II Ss’, respectively. This distinction is based on the differing semantic and syntactic properties of the involved null subjects (cf. Krapova 1998: 75). In short, while the first type of da-constructions has a pro subject, the other one contains PRO. But Krapova adds that ‘the relevant factor [responsible for this difference] is the referential (and morphological) content of embedded Tense.’ More precisely, she assumes ‘that Tense comes in two varieties – T_{nom} and T_{null}. The former corresponds to a [+T] specification and checks Nominative Case, while the latter corresponds to [-T], to indicate lack of temporal specification, and checks Null Case.’ Finally, she argues ‘that the control relation in Type II Ss is not imposed by the anaphoric properties of PRO, but follows from, or at least correlates with the specific temporal reference of the clause in which it is licensed.’ (Krapova 1998: 80)

I will adopt Krapova’s insight that Bulgarian features two varieties of da-constructions. In the next section, the hitherto discussed ingredients will be combined to arrive at an integrated proposal which leads to a more refined picture of the Bulgarian mood system, taking into account the important conclusion that da-constructions may either be subjunctive or infinitival expressions.

---

6 Two details of Siegel’s analysis should be noted: First, the attitude holder is not necessarily the speaker but rather the subject of the relevant clause. Second, the supposition of subject certainty ‘need not be old information […] [which] means that it has some assertion-like properties in that it can be used to bring new information into the discourse.’ (Siegel 2009: 1863) This proves to be of advantage when it comes to analyzing independently used da-clauses. Unfortunately, I cannot discuss this special group of da-constructions due to space limitations.
2.3 Combining both ingredients

The present proposal is a conjunction of Siegel (2009), according to whom the Balkan subjunctive is a semantic default correlated with decreased subject certainty, and of Krapova (1998) who claims that Bulgarian da-constructions may be tensed or untensed.

An important insight of this conjugation will the be that da-constructions should be characterized as subjunctives – expressing no, little or relatively decreased certainty on the part of their subject referent – only when they are finite. On the contrary, non-finite da-constructions should be characterized as infinitivals as they represent the Bulgarian equivalent (substitute) of the synchronically missing (diachronically lost) infinitive.

But before going into details, I shall briefly sketch necessary assumptions about the syntax of Bulgarian da-constructions. In (1), I used XP which still calls for a category. Krapova (1998) assumes that T in (subjunctive) da-constructions is either [+T] or [-T]. On the other hand, she locates da in C/Fin, thus analyzing it as a complementizer.\(^7\) Since I associate da with the category of mood, I will rather follow Rivero (1994, 2005) who locates da in the head of a Modal Phrase (MP) above TP.\(^8\) She adds that ‘Bulgarian resembles other Balkan languages and lacks “restructuring”/clitic climbing, which is why clitic pronouns must remain in the embedded clause/phase without exception.’ (Rivero 2005: 1086) Taking this for granted, da-constructions are always CPs.\(^9\) With this said, (1) can now be completed, thus arriving at a more detailed clause structure of Bulgarian da-constructions (possible additional projections not excluded):

\[
(1') \quad [\text{CP} \ C \ [\text{MP} \ da_M \ [\text{TP} \ [+T] \ [\text{VP} \ IND \ (VP)]]]]
\]

Now, given Krapova’s distinction of finite and non-finite da-constructions, a question that arises is how these types are distributed. The data (a part of which is cited in this paper) suggest that their distribution is as follows:

---

\(^7\) The same view is held by Penčev (1998). In her paper on the Bulgarian system of auxiliaries, Krapova (1999) proposes da to be in T, which however relates to da in periphrastic tenses, not in subjunctive clauses.

\(^8\) Rivero suggests that M is also the base position of the future particle šte, thus accounting for its complementary distribution with da. This receives additional semantic support assuming its modal origin (< present tense form of Old Bulgarian xođë̂ti ‘want’). Presumably owing to this origin, šte-clauses are always finite ([+T]).

\(^9\) This does not exclude Raising from, and Control into, da-constructions. As in many other languages, this seems to be related to the fact that they are non-indicatives which, in turn, makes them transparent for certain syntactic operations. A recent attempt to explain these transparency effects is made by Kanno & Nomura (2012).
Table 2: Distribution of non-/finite Bulgarian da-constructions.

<table>
<thead>
<tr>
<th></th>
<th>complement</th>
<th>adjunct</th>
<th>independent use</th>
</tr>
</thead>
<tbody>
<tr>
<td>finite</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>non-finite</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

Due to space limitations, I have to confine myself to the class of da-complements.

### 2.4 Da-complements

Da-complements are determined by the argument structure of the respective matrix predicate. As far as the complement position is occupied by a finite da-CP, the situation is clear: Since the da-CP is tensed, a Nominative subject is licensed. Moreover, the [+T] specification yields a ‘possible future’ (Krapova 1998: 81) interpretation of the situation denoted by the da-CP. Furthermore, as a rule such cases exhibit disjoint subject reference; cf. (2a):^10

(2) a. Iskam [ Maria da pee ]. (Smirnova 2012: 547)

  want.1SG   Maria  DA    sing.3SG

  ‘I want Maria to sing.’

b. *Iskam [ če Maria pee ].

  want.1SG   that  Maria  sing.3SG

The pair in (2) also shows that iskam ‘want’ is a predicate that is, as far as clausal complements are concerned, restricted to da-CPs. The reason is its semantic content: Iskam is a volitional verb, thus it goes well with unrealized (desired) propositions, but can by no means be sensibly combined with a complement about the truth or reality of which the subject is supposed to be certain.

Note, however, that iskam-sentences may also exhibit subject coreference as in (3):

(3) Iskaš [ da peeš ].

  want.2SG   DA    sing.2SG

  ‘You want to sing.’

---

^10 The following glosses are used: ACC = accusative; AOR = aorist; DEF = definite article; NEG = negation marker; NOM = nominative; SG = singular. Many examples are (minimal) pairs, contrasting da-constructions with če-clauses. For the sake of consistency, the da-example is always given first. Sometimes, this required switching the original sequence.
Assuming the existence of non-finite da-CPs, it seems natural that the da-CP in (3) is non-finite.\footnote{An analysis arguing for finite-only complements in Bulgarian would have to propose that the da-CP in (3) contains a pro subject which is interpreted as being coreferent with the matrix pro subject.} Still, there are two possible analyses for iskam: It may either be a Raising or a Control predicate. In the former case, the subject of the da-CP moves into the matrix structure, leaving a trace in its base (and intermediate) position(s). In the latter case, the da-CP contains PRO which is controlled by the subject of iskaš. Both cases are schematized in (4):

\begin{align*}
(4) & \quad \text{a. } \text{DP}_1 V [\text{CP} \dots t_1 \text{ da}_M T_{[-T]} \dots ] \quad \text{(Raising from a non-finite da-CP)} \\
& \quad \text{b. } \text{DP}_1 V [\text{CP} \dots \text{PRO; da}_M T_{[-T]} \dots ] \quad \text{(Control into a non-finite da-CP)}
\end{align*}

I leave open the question which of these analyses is more adequate to capture (3). What is more important is that non-finite da-constructions may generally occur in these two environments, depending on the properties of their relevant matrix predicate. Thus, while moga ‘can’ behaves like a Raising predicate, započvam ‘begin’, zabravjam ‘forget’ and znam ‘know how’ behave like Control predicates.\footnote{Krapova (1998: 74) subsumes these verbs under the term control verbs, i.e. verbs which take in any non-freely interpreted empty category (cf. Joseph 1992). She also mentions opitvam se ‘try’ and uspjivam ‘succeed’ and notes that some of these verbs show ambiguities in their behavior as Raising rather than Control predicates.} Noticeably, Rivero remarks that ‘Bulgarian auxiliaries resemble Balkan raising/control verbs, so take full clauses as complements.’ (Rivero 2005: 1090) This supports the above-mentioned claims, namely that non-finite da-constructions are CPs and that they may occur in Raising or Control structures. Importantly, it also means that the relevant matrix predicates either are auxiliaries or assume auxiliary-like properties when combined with non-finite da-constructions.

A Raising case is illustrated in (5). It features the modal verb moga ‘can’. It is a Raising predicate as it provides an external argument position to which it does not assign a thematic role. As a consequence, the subject from within the non-finite da-CP can be raised into the matrix where the Nominative is licensed. Notice that the bracketting in (5) indicates that two distinctly realized subject DPs are ruled out (cf. Lempp 1981: 62), thus corroborating the view that the involved da-CP is non-finite:

\begin{align*}
(5) & \quad \{\text{az}\} \text{ moga } \{\text{az}\} \text{ da rabotja (cf. Lempp 1981: 62)} \\
& \quad \text{I can.1SG I DA work.1SG} \\
& \quad \text{‘I can work’}
\end{align*}
The Control case is illustrated in (6). Since znam ‘know (how)’ assigns a thematic role to its external argument position, the embedded subject cannot be raised. It follows that (6) must contain PRO. As noted by Krapova (1998: 75), the latter does not alternate with lexical DPs. Much like in (5), this means that two distinctly realized subject DPs are ruled out.

(6) Znam [ da peja ].  
    want.1SG DA sing.1SG  
    ‘I know how to sing.’

While znam in (6) shows a modal interpretation (‘know how’; Krapova 1998: 74), it is interpreted veridically (‘know’) when combined with an indicative če-clause; cf. (7b):

(7) a. *Znam [ Maria da peć ].  
    know.1SG Maria DA sing.3SG
b.  Znam [ če Maria peć ].  
    want.1SG that Maria sing.3SG  
    ‘I know that Maria sings.’

Under the desirable assumption that there is only one lexical entry for znam, its modal and veridical interpretations, respectively, should arise due to the properties of its complement. Accepting Siegel’s (2009) subjunctive-as-default analysis, the veridical interpretation results from the amalgamation of subject certainty (indicative) with the lexical content KNOW. But why is znam restricted to occur with non-finite da-CPs as in (6), while it cannot be used with finite da-CPs as in (7a)? I suggest the following explanation: Due to their tense specification, finite da-CPs are interpreted as hypothetical/unrealized situations (cf. Krapova 1998: 82).  

However, this kind of situation is completely incompatible with KNOW: Since knowing something clearly requires to be certain about it, finite da-CPs are ruled out. On the other hand, non-finite da-CPs are not necessarily interpreted as hypothetical/unrealized situations due to

---

13 In Siegel’s (2009) terms, they are correlated with subject uncertainty due to their semantic vacuity. At this point, it may seem that the interpretations suggested by Krapova (1998) and Siegel (2009) contradict each other. Note, however, that Siegel’s notion of subject uncertainty relates to the category of mood, while the interpretations suggested by Krapova follow from properties of embedded tense. Assuming that mood and tense interact, the contradiction turns into a plausible explanation: It is exactly due to the subjunctive – marking the absence of subject certainty – that the embedded tense receives a hypothetically, unrealized etc. interpretation.
the fact that they are untensed. I suggest that they are interpreted as bare events instead, and conclude that combining such a bare (abstract) event with \textsc{know} yields the modal interpretation observable in (6).

Bare events also play a role in the interpretation of sentences featuring matrix verbs of perception; cf. (8):

\begin{verbatim}
(8) Vidjaj (ja) [ Mila da zatvarja prozoreca ]. (cf. Werkmann 2007: 3)
   see.AOR.1SG her.ACC Mila.ACC DA close.3SG window.DEF
   ‘I saw Mila close the window.’
\end{verbatim}

Werkmann (2007) observes that (8) expresses a \textit{direct perception report}. Note that the embedded subject \textit{Mila} is in the Accusative, which is indicative of \textit{Exceptional Case Marking} (ECM). Within the current proposal, these facts can be explained as follows: Higginbotham (1986) claims that direct perception amounts to perceiving bare events. Assuming that non-finite \textit{da}-CPs may denote bare events, the fact that (8) is a \textit{direct perception report} follows straightforwardly as the perception of bare events entails the perceived event to be included in the perceiving situation – it takes place before the eyes of the perceiver. As concerns ECM, the first thing to mention is that the Nominative on the embedded subject is not licensed due to \([-T]\). Furthermore, Raising is no option for the embedded subject to receive case as both argument positions of the matrix verb are already saturated. Control is no option either, since the \textit{da}-CP occupies the direct object position of the verb of perception, so that \textit{Ivo} cannot be claimed to occupy this matrix position from where it might control a potential PRO within the \textit{da}-CP.\footnote{Werkmann (2007) suggests a different analysis: The embedded subject is raised into the matrix object position where it receives a thematic role and the Accusative. Note that this implies the assignment of two thematic roles to one DP (and in fact also two matrix object positions), which is why I consider an ECM analysis more suitable.} It follows that the only way of saving the structure – and of expressing a \textit{direct perception report} in Bulgarian – is ECM: The embedded subject is exceptionally assigned the Accusative by the matrix predicate.

But Bulgarian verbs of perception may also combine with other types of complements:

\begin{verbatim}
(9) a. Viždam [ Maria[/tja] da idva ]. (Smirnova 2012: 553)
   see.1.SG Maria/she.NOM DA come.3SG
   ‘I see Maria coming.’
\end{verbatim}

\footnote{Importantly, non-finite \textit{da}-CPs are not restricted to denote bare events. They may also be interpreted as hypothetical, unrealized etc. situations (propositions) in appropriate contexts. On the contrary, finite \textit{da}-CPs are restricted to exactly this kind of interpretation due to their tense specification.}
Following Smirnova, uttering (9a) is felicitous in a context where the perceiver sees a woman approaching, but is not entirely sure that it really is Maria. This interpretation differs crucially from the direct perception interpretation of (8). Rather, it is reminiscent of Siegel’s claim that subjunctives express *subject uncertainty*. The semantic difference between (8) and (9a) can be accounted for assuming that these sentences differ in the finiteness of their *da*-CPs: While (8) contains a non-finite *da*-CP, the one in (9a) is finite. Hence, the interpretation of (9a) arises on the grounds of *subject uncertainty*, while the interpretation of (8) follows from the fact that non-finite *da*-CPs have the potential of denoting bare events, thus enabling the expression of *direct perception reports* (Higginbotham 1983).

Finally, verbs of perception may also combine with indicative CPs as shown in (9b). Uttering (9b) is felicitous when ‘the subject of perceptual experience is absolutely certain that the proposition \( p \) denoted by the embedded clause is true’ (Smirnova 2012: 552). This description is nearly identical with Siegel’s (2009) *subject certainty*. Therefore, it is legitimate to assume that the ‘certainty interpretation’ is due to the indicative. On the other hand, Werkmann (2007) notes that sentences like (9b) may be interpreted as *indirect perception reports*. This, in turn, is due to the fact that \( \text{če} \)-clauses denote propositions (independent situations), not bare events as non-finite *da*-CPs. I already mentioned that the perception of a bare event implies this event to be included in the perceiving situation. Such an inclusion is, however, not entailed in the case of perceiving an independent situation. In (9b), for example, it is absolutely possible to obtain a direct perception interpretation, but only if Ivo’s smoking takes place in front of the perceiving individual. Importantly, however, it may also be the case that the perceiver merely infers that Ivo smokes based on certain details in the situation that he or she is actually perceiving (such as, e.g., smoke, ashes, a burning cigarette or the like).

I will finish my illustrations with the minimal pair cited in (10). It features the matrix verb *mislja* ‘think’. This time, the relevant opposition is affirmative vs. negated:

\[
\text{(10) a. Mislja [ \{če / *da\} Paulina e izjala tortata ]. (Siegel 2009: 729)} \\
\text{think.1SG \ that \ DA Paulina be.3SG eaten cake.DEF} \\
\text{‘I think that Paulina ate the cake.’}
\]
Interestingly, a *da*-complement is ruled out in the affirmative example (10a), but becomes available under sentential negation as in (10b). Since sentence polarity seems to be the factor determining the un-/availability of the finite *da*-construction, this case may be classified as an instance of so-called *polarity subjunctive*. Siegel argues that indicative is associated with a greater degree of subject certainty. When applied to the negation cases, this means that indicative is associated with the meaning that the subject believes in the negation of the embedded clause. That is, indicative gets what has been called the Neg-Raising/strengthening reading” (Siegel 2009: 1874) Hence again, the subject-as-default analysis works. Moreover, this is another example revealing that the form of a complement may affect the interpretation of a whole utterance. More specifically, the choice of one or another type of complement serves to disambiguate differing interpretations in Bulgarian.

4. Summary

The goal of this paper was to account for the interpretation of Bulgarian *da*-constructions, especially complements. In order to achieve this goal, I have combined two ingredients: First, I assumed Siegel’s (2009) view according to which the indicative is associated with the supposition of *subject certainty*, while the subjunctive is a semantic default used whenever the indicative supposition is too specific. But as Bulgarian lacks a morphological infinitive, not all *da*-constructions are subjunctives correlated with *subject uncertainty*. What I tried to show is that *da*-CPs may also function as equivalents of infinitival structures in other language. I propose the term *non-indicative* to capture both functions of *da*-constructions. With this said, Table 1 from above can be augmented as follows:

<table>
<thead>
<tr>
<th>Mood</th>
<th>finite</th>
<th>mood</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\emptyset$</td>
<td>+</td>
<td>indicative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(subject certainty)</td>
</tr>
<tr>
<td><em>da</em></td>
<td>+</td>
<td>non-indicative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(→)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>infinitival</td>
</tr>
</tbody>
</table>
In sum, it was shown that Bulgarian is likely to possess non-finite complements, namely non-finite *da*-CPs. Special evidence is supplied by the intricate class of verbs of perception, which may take three different types of complements, yielding three different interpretations. Furthermore, I argued *da* to be a modal (mood) particle not associated with the notion of finiteness. As a result, there is no need to assume more than one lexical entry for *da*. Moreover, *da* is semantically vacuous and, thus, a mere morphosyntactic placeholder marking what I call the Bulgarian *non-indicative*. Among other things, future work will have to examine whether the present assumptions apply to the whole range of relevant Bulgarian examples, and whether they extend to *da*-adjuncts as well as to independently used *da*-clauses.

References
