

THE SYNTAX OF FACTIVITY AND REFERENTIALITY IN SOUTH SLAVIC LANGUAGES *

Tomislav Sočanac (socanact@gmail.com)

1. INTRODUCTION: FACTIVITY AND PRESUPPOSITION

The paper studies the linguistic manifestations of factivity and a number of related notions, such as presuppositionality and referentiality (the latter being viewed as synonymous to discourse givenness or familiarity). The focus of the paper is the factive and referential predicates and their complements in South Slavic languages, specifically Bulgarian (Bg) and Bosnian/Croatian/Montenegrin/Serbian (BCMS). The main goal of the paper is to classify different types of verbs in relation to the notions of factivity and referentiality (both of which will be defined more precisely in the subsequent parts of the paper) and then provide a syntactic analysis that accounts for the contrasting properties observed with the relevant verb types in this context.

Ever since the seminal work in Kiparsky & Kiparsky (1970), there has been a rich and growing literature dealing with the properties of factive predicates and their complements.¹ Below, we give a (non-exhaustive) list of predicates (taken from Anand et al. [2017]) that roughly encompass the class of factive-type verbs (as they have been traditionally labeled in the literature):

(1) ‘Factive’ verbs

absorb, acknowledge, ascertain, be aware/conscious, be happy/sad, be surprised, bother, catch on, comprehend, detect, discover, figure out, find out, forget, grasp, hear, intuit, know, learn, notice, observe, prove, realize, recall, recollect, register, resent, regret, remember, see, take in, understand.

What distinguishes factive-type predicates such as those in (1) from non-factives is that they introduce propositions which are typically assumed to be true by the relevant discourse participants (e.g. the speaker and the subject), which is not the case with non-factive verbs. As a result, complements introduced under non-factive predicates (e.g. *think*, *believe*, *say*, etc.) can be contradicted by the speaker (as shown in [2]), whereas those introduced under factive-type verbs typically cannot be contradicted in this way (3). The contrast between factives and non-

* Some of the research results presented in this paper were enabled by funding received from the European Union’s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No.898425. I would also like to extend a special acknowledgement to Iliyana Krapova, since the present paper would not have been possible without the insights gleaned from our extensive collaboration over the last several years. I am also thankful to Guglielmo Cinque, Roland Hinterhölzl, Lena Baunaz, as well as the audiences of the workshop on *Mood and Modality Markers Across Languages* (Venice, 2024), *Biennial conference on Balkan and South Slavic Linguistics, Literature and Folklore* (Columbus OH, 2022) and *Slavic Linguistic Society 14* (Potsdam, 2019), for their helpful comments and suggestions. Any possible errors or insufficiencies contained in the paper are my own.

¹ For more theoretical background on factives, see also Karttunen (1971), Haegeman (2006), de Cuba (2007), Kastner (2015), Djärv (2019), etc.

factives in this context has been noted on a wide cross-linguistic basis, as illustrated below on the examples of English, French (Fr), Bg and BCMS.

(2) a. John believes that Mary left, but she is still here.
 b. Jean affirme que Marie est partie, mais elle est toujours là. (Fr)
 J. claims that M. has left, but she is still here
 c. Ivan kazva če Marija e zaminala, no tja ošte e tuk. (Bg)
 I. says that M. has left but she still is here
 d. Ivan misli da je Marija otišla, ali ona je još tu. (BCMS)
 I. thinks that has M. left but she is still here

(3) a. # John knows that Mary left, but she is still here.
 b. # Jean est content que Marie soit partie, mais elle est toujours là. (Fr)
 J. is happy that M. has left, but she is still here
 c. # Ivan sǎžaljava, če Marija e zaminala, no tja ošte e tuk. (Bg)
 I. regrets that M. has left but she still is here
 d. # Ivan zna da je Marija otišla, ali ona je još tu. (BCMS)
 I. knows that has M. left but she is still here

The factive pattern in (3) has usually been explained with reference to the notion of presuppositionality: factive predicates like *know* or *regret* introduce a presupposition which is taken to be true by the discourse participants, including the speaker, and hence their complements cannot be felicitously contradicted by the speaker, unlike those of non-factive predicates in (2) (Kiparsky & Kiparsky 1970, Karttunen 1971, 1974, etc.).

The linguistic literature on presuppositions is vast and fraught with disagreements, so any kind of representative literature review in this context is not practical due to space constraints.² I will focus instead on those properties of presuppositional clauses that will be relevant for the argument presented in the paper, and that will help me classify different types of South Slavic verbs in relation to factivity and referentiality. Section 2 deals with the notion of ‘presupposition projection’ (a term first put forward in Langendoen & Savin 1971), which I define here as the property of factive predicates to preserve the truth reading of their complement even when they are embedded under non-veridical operators such as negation, questions, or modals. Section 3 discusses the contrast between factive and non-factive predicates in relation to referentiality or discourse givenness, showing that factive clauses typically only convey old information that is already grounded in discourse. Section 4 then uses the insights from Sections 2 and 3 to develop a fine-grained classification of different predicates in relation to factivity and referentiality. Section 5, which constitutes the theoretical core of the paper, analyzes how the semantic contrasts between factives and other types of verbs discussed in Section 4 are encoded at the level of syntactic structure. The broader point that will be made there is that both factivity and referentiality are syntax-semantics interface phenomena encoded within the clausal structure, and thus not reducible just to the lexical semantics of different predicates. Finally, Section 6 summarizes and concludes the paper.

² See Karttunen (1974), Stalnaker (1974), Heim (1992), Abbott (2000) or Abrusán (2011) (among many others) for a more detailed account of the linguistic properties of presuppositional clauses and the theoretical controversies related to this topic.

2. PRESUPPOSITION PROJECTION: FACTIVE VS. VERIDICAL VERBS

Besides the contrast between factive and non-factive predicates previously noted in (2)-(3), pertaining to the cancellability of the truth reading of the embedded complement, there are a number of additional properties whereby factive presuppositional clauses differ from non-factive ones. One of those is the property of ‘presupposition projection’, which refers to the fact that the truth reading associated with factive/presuppositional predicates is maintained even when they are embedded under non-veridical operators such as negation, questions or modals.³

- (4) a. John does not [know/regret] that Mary left: # she is still here.
- b. Does John [know/regret] that Mary left? # She is still here.
- c. John may [know/regret] that Mary left, # but she is still here.
(>>presupposition: ‘Mary left’)⁴

- (5) a. Ivan ne [znae / sūžaljava], če Marija e zaminala: # tja ošte e tuk. (Bg)
I. NEG knows regrets that M. has left she still is here
- b. Ivan [znae / sūžaljava] li, če Marija e zaminala? # Tja ošte e tuk.
I. knows regrets Q that M. has left she still is here
- c. Ivan može by [znae / sūžaljava], če Marija e zaminala, # no tja ošte e tuk.
I. maybe knows regrets that M. has left but she still is here

- (6) a. Ivan ne [zna / žali] da je Marija otišla: # ona je još tu. (BCMS)
I. NEG knows regrets that has M. left she is still here
- b. [Zna / žali] li Ivan da je Marija otišla? # Ona je još tu.
knows regrets Q I. that has M. left she is still here
- c. Ivan možda [zna / žali] da je Marija otišla, # ali ona je još tu.
I. maybe knows regrets that has M. left but she is still here

As the examples in (4)-(6) show, the speaker cannot contradict the truth presuppositions associated with the complements embedded under factive-type verbs even when the latter are used in non-declarative and non-veridical contexts. This is what the term ‘presupposition projection’ refers to in the present paper: in effect, the presuppositional reading associated with factives projects outside the scope of the non-veridical operators such as negation, questions or modals.

The presuppositionality tests in (4)-(6) also allow us to distinguish factive-type predicates from other verbs associated with a strong truth reading, in particular veridical verbs. According to the definition of veridicality proposed in Egré (2008) (which I will rely on here as well), veridical readings are based on the semantic relation of truth entailment.⁵ In simple declarative contexts, veridicality/truth entailment functions similarly as presuppositions in factive contexts, in the sense that the truth reading of the clause is not cancellable by the speaker.

³ I refer to these operators as non-veridical because they cancel out the semantic truth-entailment relationship that characterizes veridical-type clauses (according to the definition of veridicality proposed in Egré [2008], which I am adopting here as well). I will discuss the notions of (non)veridicality and truth entailment later in this section.

⁴ Presuppositions will be marked with the symbol >> (a standard annotation for presuppositions in the literature), while the absence of presuppositional readings (where relevant) will be marked with the symbol //>>.

⁵ Truth entailment will be marked with the symbol \models (once again, a standard annotation in this context). The definition of truth entailment assumed here is based on a simple semantic/logical relation of $Vp \models p$, i.e. if the sentence as a whole is true, then the embedded proposition is necessarily true as well. The contrast between veridical (i) and non-veridical (ii) predicates in this context is illustrated below.

(i) It is true that Mary left. \models Mary left (truth entailment)
(ii) It is possible that Mary left. // \models Mary did not necessarily leave (no truth entailment)

This is demonstrated in (7) on the example of veridical predicates such as ‘it is true’ and ‘it is clear’.

(7) a. It is [true/clear] that Mary left, # but she is still here.
 b. [Vjarno/jasno] e, če Marija e zaminala, # no tja ošte e tuk. (Bg)
 true clear is that M. has left but she still is here
 c. [Istina/jasno] je da je Marija otišla, # ali ona je još tu. (BCMS)
 true clear is that has M. left but she is still here

Nevertheless, the main difference between truth entailment and presuppositionality is that the former, unlike the latter, breaks down in the presence of non-veridical operators such as negation, as shown in (8) below.⁶

(8) a. It is not [true/clear] that Mary left: I think she is still here.
 b. Ne e [vjarno/jasno], če Marija e zaminala: mislja, če ošte e tuk. (Bg)
 NEG is true clear that M. has left I-think that still is here
 c. Nije [istina/jasno] da je Marija otišla: mislim da je još tu. (BCMS)
 NEG-is true clear that has M. left I-think that is still here

The paper will focus on factive/presuppositional predicates (among others), while leaving veridical predicates to the side.

3. FACTIVE VERBS AND REFERENTIALITY: OLD VS. NEW INFORMATION

Another well-known cross-linguistic property of factive/presuppositional predicates is the fact that their complements typically cannot be used to introduce new information which updates the common ground, i.e. the shared knowledge between discourse participants.⁷ Note the contrast between (9) and (10) in this context:

(9) a. I [think/believe] that Mary left.
 b. [Vjarjam/mislja], če Marija e zaminala. (Bg)
 I-believe I-think that M. has left
 c. [Vjerujem/mislim] da je Marija otišla. (BCMS)
 I-believe I-think that has M. left

(10) a. I [regret/am surprised] that Mary left.
 b. [Sūžaljavam / iznenadan sūm], če Marija e zaminala. (Bg)
 I-regret surprised I-am that M. has left
 c. [Žalim / iznenaden sam] da je Marija otišla. (BCMS)
 I-regret surprised I-am that has M. left

The parts of speech underlined in (9)-(10) constitute what Simons (2007) has dubbed the ‘main point of the utterance’ (MPU), i.e. the most informative part of the sentence that updates the common ground with new information. In the case of non-factive verbs such as *think* or *believe* in (9), the MPU is the proposition contained in the embedded complement (i.e. ‘Mary left’). With factive verbs, however, the MPU is not the information contained in the embedded clause,

⁶ I only demonstrate the breakdown of truth entailment on the examples of negated predicates (due to space constraints), but the same result obtains in questions and modal contexts as well.

⁷ For more on the notion of common ground, see Stalnaker (2002) and the references therein.

given that the latter introduces a presupposition which is assumed to be known by the addressee. The new information in this case pertains to the emotional attitude or reaction of the speaker (e.g. regret, surprise) in relation to the propositional content of the embedded clause.

A more precise diagnostic that can be used to demonstrate the contrast between factive and non-factive verbs in relation to MPU are question and answer pairs. Given that the main function of an answer to a question is to provide new information previously unknown to the addressee, non-factive verbs can be used in this context (11a-c) whereas factives typically cannot (11a'-c').⁸

(11) When does the game begin?

- a. I [think/believe] it begins at 7 o'clock.
- a.' # I [regret/am surprised] that it begins at 7 o'clock.
- b. [Mislja / vjarvjam], če započva v 7 časa. (Bg)
I-think I-believe that begins at 7 o'clock
- b.' # [Sūžaljavam / iznenadan sūm], če započva v 7 časa.
I-regret surprised I-am that begins at 7 o'clock
- c. [Misljam/vjerujem] da počinje u 7 sati. (BCMS)
I-think I-believe that begins at 7 o'clock
- c.' # [Žalim / iznenađen sam] da počinje u 7 sati.
I-regret surprised I-am that begins at 7 o'clock

The property of factive verbs to introduce old, discourse-linked information will be subsumed under the label 'referentiality' (a term borrowed from de Cuba [2007, 2017] and de Cuba & Ürögdi [2010], among others, but used in a different sense here⁹).

Note, however, that not all referential predicates are necessarily factive, because there are certain verbs which refer back to old information but do not presuppose that the information in question is true. This is the case, for instance, of negative epistemic or assertive verbs such as *doubt* or *deny*. Such verbs cannot be used in an answer to a question (12), given their referential properties, but their complements can be contradicted by the speaker because they do not introduce presuppositions (13).

(12) When does the game begin?

- a. # I [doubt/deny] that it begins at 7 o'clock.
- b. # [Sūmnjavam se/otričam], če započva v 7 časa. (Bg)
I-doubt I-deny that begins at 7 o'clock
- c. # [Sumnjam/poričem] da počinje u 7 sati. (BCMS)
I-doubt I-deny that begins at 7 o'clock

(13) a. John [doubts/denies] that Mary committed the crime, and I believe she is innocent too.
//>> no presupposition: Mary did not (necessarily) commit the crime.

⁸ The examples below (and similar examples recurring throughout the paper) are taken and adapted from Djärv, Heycock & Rohde (2017).

⁹ I view referentiality as synonymous with discourse givenness or familiarity, which is slightly different from the way in which the authors listed above defined this notion. I also do not adopt the syntactic analysis proposed by de Cuba (2007) in this context. He argued that referential clauses contain a smaller syntactic structure than non-referential clauses (lacking the CP projection in particular), while I will argue that referential clauses project a larger structure, with the nominal DP on top of the clausal CP (see Section 5 for more detail).

- b. Ivan [se sūmnjava/otriča], če Marija e izvürshila prestüplenieto, i az sūshто I. doubts denies that M. has committed crime and I also mislja, če tja e nevinna. think that she is innocent (Bg)
- c. Ivan [sumnja/poriče] da je Marija izvrsila zločin, i ja također mislim da je I. doubts denies that has M. committed crime and I also think that is nevina. innocent (BCMS)

To summarize, we have so far identified at least 3 distinct groups of predicates in relation to factivity and referentiality: (i) non-factive/non-referential predicates (e.g. *think*, *believe*, *say*, etc.), which do not involve presuppositional readings (2) and can be used to update the common ground (9), (11a-c); (ii) factive/referential predicates (e.g. *regret*, *be happy*, *know*, etc.), which typically project presuppositions (3)-(6) and are not used to update the common ground (11a'-c'); (iii) non-factive referential predicates (e.g. *doubt*, *deny*, *disprove*, etc.), which pattern with verbs in (i) in that they do not involve truth-presuppositional readings (13) but also pattern with those in (ii) in that they can only bring forth old information known from the previous discourse (12). Nevertheless, this classification will require further refinement because, as we will see in the following section, not all factive-type verbs listed in (1) exhibit identical properties in relation to factivity and referentiality.

4. VERB CLASSIFICATION

It has been noted in the literature for some time (at least since Karttunen [1971]) that not all factive-type predicates exhibit identical properties in relation to factivity and presuppositionality. In order to illustrate the relevant contrasts in this context, let us return once again to the property of presupposition projection, i.e. the preservation of truth-presuppositional readings under non-veridical operators. Even though factive-type predicates such as those previously listed in (1) tend to preserve their presuppositional status when appearing under negation or in questions, they exhibit some contrasts when they are associated with other types of non-veridical operators, in particular conditionals (Karttunen 1971, 1973). Note the difference between (14a) and (14b).¹⁰

- (14) a. If I later [realize/find out/learn] that I was wrong, I will apologize (but right now, I believe that I am right).
//> no presupposition, i.e. the speaker was not (necessarily) wrong
- b. If I later [regret/become sorry] that I was wrong, I will apologize (# but for now, I believe that I am right).
>> presupposition: the speaker was wrong

Predicates featured in (14a) lose their presuppositional status when appearing in conditional clauses of the type exemplified above, unlike those in (14b) which maintain their presuppositional reading in this context as well. If we look more closely at the semantic properties of the predicates featured above, we can notice that those in (14a) are verbs of a more cognitive type, used to describe how the speaker attained (or will attain) the knowledge of a given state of affair (e.g. *realize*, *find out*, *learn*, etc.), whereas those in (14b) are of a more

¹⁰ The examples are taken from Karttunen (1971) and adapted for the purposes of the present paper.

emotive type, describing the speaker's emotional attitude or reaction towards a given state of affairs (e.g. *regret*, *be sorry*, etc.).

The same type of contrast we noted in (14) in the case of English obtains in the South Slavic languages under study as well:¹¹

(15) a. Ako edin den [osúznaja / razbera], če súm ja obidil, šte se izvinja (no if one day I-learn I-realize that I-have her insulted will apologize but zasega mislja, če ne súm ja obidil).
for-now I-think that NEG I-have her insulted
'If one day I [learn/realize] that I have insulted her, I will apologize (but for now I don't think that I insulted her).'

b. Ako edin den súžalja, če sám ja obidil, šte se izvinja (# no zasega mislja, if one day I-regret that I-have her insulted will apologize but for-now I-think če ne súm ja obidil).
that NEG I-have her insulted
'If one day I regret that I have insulted her, I will apologize (# but for now I don't think that I insulted her).'

(16) a. Ako kasnije [shvatim/doznam] da sam ju uvrijedio, ispričat ču se (ali if later I-realize I-find out that I-have her insulted apologize I-will but zasad mislim da ju nisam uvrijedio). (Bg)
for-now I-think that her not-have insulted

b. Ako kasnije zažalim da sam ju uvrijedio, ispričat ču se, (# ali zasad if later I-regret that I-have her insulted apologize I-will but for-now mislim da ju nisam uvrijedio). (BCMS)
I-think that her NEG-have insulted

Once again, factive emotive verbs maintain their presuppositional status when appearing in conditionals in South Slavic languages as well (15b)/(16b), whereas cognitive factives do not (15a)/(16a).

A similar contrast between emotive and cognitive factives is also observed in relation to referentiality, i.e. introducing old vs. new information. Cognitive factives are not always referential in this sense, because they can sometimes be used to update the common ground with new information (17a-c), whereas this is not the case with emotive factives (17a'-c'). Note the contrasts below related to the possibility of using factive-type verbs in answers to questions (Simons 2007, Djärv et al. 2017).

(17) When does the game begin?

a. I just [realized/found out/learned] that it begins at 7.
a.' # I [regret/am surprised] that it starts at 7.
b. [Osúznah / otkrih], če započva v 7. (Bg)
I-learned I-found out that begins at 7

b.' # [Súžaljavam/iznenadan súm], če započva v 7.
I-regret surprised I-am that begins at 7

c. [Doznah / otkrih] da počinje u 7. (BCMS)
I-learned I-found out that begins at 7

c.' # [Žalim / iznenađen sam] da počinje u 7.
I-regret surprised I-am that begins at 7

¹¹ I am grateful to Iliyana Krapova for providing me with the Bg examples in (15).

The contrasts we noted in (14)-(17) are some of the reasons why cognitive factive verbs have usually been treated differently than emotive factives in the literature (going all the way back to Kiparsky & Kiparsky [1970] or Karttunen [1971], among others). Given that cognitive factives exhibit a number of properties which are not typical of factive/presuppositional clauses, they have often been described as ‘semi-factive’ verbs in the literature (e.g. Karttunen 1971, 1973). Emotive factives, on the other hand, can be considered as genuine factive verbs. I will adopt the same labelling in this paper as well.

At this point, therefore, we can distinguish between (at least) four different verb groups that exhibit contrasting properties in relation to factivity and referentiality. The verb class previously grouped in (i), containing assertive verbs (e.g. *say*, *claim*, *affirm*, etc.) and propositional-attitude verbs (e.g. *think*, *believe*, *consider*, etc.), will be subsumed under the label ‘volunteer-stance verbs’, first proposed in Cattell (1978), because they express speech acts and thus involve an active contribution of the speaker to the common ground. As noted before, these types of verbs are neither factive nor referential. The second group (listed below in [ii]) involves referential predicates which cannot introduce new information but which also do not involve truth-presuppositional readings (e.g. *doubt*, *deny*, *disprove*, etc.). The third group contains predicates which were just dubbed semi-factives, given that they tend to involve factive readings by default but there are a number of contexts (some of which we have observed here as well) in which they lose their factive/presuppositional status (e.g. *find out*, *discover*, *realize*, *learn*, *prove*, etc.).¹² Finally, the last group contains emotive factive predicates, which I will consider here as ‘true factives’ (e.g. *regret*, *be glad*, *be surprised*, etc.). The relevant verb groups are listed below:

(18) **Verb classification in relation to factivity and referentiality**

- (i) Volunteer-stance verbs (e.g. *say*, *claim*, *think*, *believe*, etc.).
- (ii) Non-factive referential verbs (e.g. *doubt*, *deny*, *disprove*, *debunk*, etc.)
- (iii) Semi-factive verbs (e.g. *realize*, *find out*, *discover*, *learn*, *prove*, etc.)
- (iv) True factive verbs (e.g. *regret*, *lament*, *be glad*, *be happy*, *be surprised*, etc.)

As we will see in the following section, each of the verb groups listed in (i)-(iv) above is associated with a different type of underlying syntactic structure, which, in turn, explains the contrasting properties that they were shown to exhibit.

5. FORMAL ANALYSIS

I will argue that the contrasts we noted between different types of predicates in relation to factivity and referentiality are encoded at the level of the syntactic structure of their clausal complements, specifically at the left periphery of the clause. As we will see in more detail shortly, factivity is encoded at the level of CP while referentiality is encoded at the level of DP (the latter corresponding to an additional nominal projection heading the clausal CP). I will argue that volunteer-stance verbs in (i) involve CP structures while non-factive referential verbs (ii) and true factives (iv) involve DP structures. Semi factives in (iii), on the other hand, may project either a CP or a DP structure, which can explain the mixed properties in relation to

¹² The cognitive factive verb *know* is more difficult to neatly classify in this context, because it seems to have a more robust presuppositional reading than other cognitive semi-factives, but there are also some environments in which its presuppositional status breaks down. For instance, it loses its factive reading when appearing in a construction such as ‘as far as I know’ (Karttunen 1971). A more precise analysis of the verb *know* in relation to factivity and referentiality will be left for future work.

factivity and referentiality that they were shown to exhibit in Section 4. I will demonstrate these structural differences via a number of relevant syntactic tests and diagnostics.

Let us first focus on the contrasts between CP and DP structures, which are relevant in the context of (non)referentiality. The simplest diagnostic that can be used in order to demonstrate whether the embedded clause corresponds to a CP or a DP structure is to try to insert an additional (pro)nominal element on top of the clausal CP hosting the complementizer. This was the diagnostic originally used in Kiparsky & Kiparsky (1970) in order to demonstrate the special syntactic status of factive verbs and their complements. Note the grammaticality contrast in (19):

(19) a. I regret (the fact) that Mary left.
 b. I [think/say] [*the fact/*the claim/*the opinion] that Mary left.

Factive verbs such as *regret* allow for the optional insertion of a nominal constituent (i.e. *the fact*) heading their embedded complement, which indicates that they introduce an additional nominal DP layer on top of the clausal CP.

A similar pattern is observed in South Slavic languages as well, the only difference being that these languages do not introduce a full nominal constituent such as ‘the fact’ on top of the embedded CP, but use light pronominal items instead: *tova* ‘it, that’ in Bg and *to* ‘it, that’ in BCMS. The verb groups previously listed in (ii)-(iv) allow for the (optional) insertion of these pronominal items on top of the embedded CP; only volunteer-stance verbs in (i) do not allow for this option. The relevant patterns for each verb group from (i)-(iv) are given in the South Slavic examples below:

(20) a. [Mislja / struva mi se] (*tova), če Marija e zaminala. (Bg)
 I-think seems to-me it that M. has left
 b. [Misljam / čini mi se] (*to) da je Marija otišla. (BCMS)
 I-think seems to-me it that has M. left
 ‘[I think/it seems to me] (*it) that Marija left.’

(21) a. [Sūmnjavam se/otričam] (tova), če Marija e zaminala. (Bg)
 I-doubt I-denry it that M. has left
 b. [Sumnjam/poričem] (to) da je Marija otišla. (BCMS)
 I-doubt I-denry it that has M. left
 ‘I [doubt/deny] (it) that Marija left.’

(22) a. [Razbrah / otkrih] (tova), če Marija e zaminala. (Bg)
 I-realized I-found out it that M. has left
 b. [Shvatih / otkrih] (to) da je Marija otišla. (BCMS)
 I-realized I-found out it that has M. left
 ‘I [realized/found out] (it) that Marija left.’

(23) a. Sūžaljavam (za tova) če Marija e zaminala.¹³ (Bg)
 I-regret for it that M. has left

¹³ Factive-emotive verbs such as those in (23) exhibit some special properties in this context as well, because they tend to introduce an additional preposition on top of the nominal DP (as in the Bg example in [23a], for instance). As a result, Krapova (2010) argued that factive emotive verbs select oblique PP complements in general. This issue, however, is not crucial for the present discussion: what counts is that complements to factive-emotive verbs always contain a DP projection as well (regardless of whether the latter is headed by an additional PP projection or not).

b. Žalim (to) da je Marija otišla. (BCMS)
 I-regret it that has M. left
 'I regret (it) that Marija left.'

The grammaticality contrasts above are the first indication that volunteer stance verbs in (i) project a 'smaller' embedded CP structure with no nominal layer on top, whereas all other verb groups (ii)-(iv) allow for a DP layer above CP.

Another useful diagnostic in this context is the possibility to front the complement clause to the matrix subject condition, which was shown to be a property of DP complements (given the standard minimalist assumption that only DP constituents can check the uninterpretable D feature in T). As expected, therefore, this operation is not allowed with volunteer stance verbs (24), while it is possible with other verb types discussed above (25).¹⁴

(24) a. * (* Tova) će Marija e zaminala, [mislja / struva mi se]. (Bg)
 it that M. has left I-think seems to-me
 b. * (* To) da je Marija otišla, [mislim/mi se čini]. (BCMS)
 it that has M. left I-think to-me seems
 * 'That Marija left, [I think/it seems to me].'

(25) a. (Tova) će Marija e zaminala, az [znam/ otričam/veče razbrah]. (Bg)
 it that M. has left I know deny already realized
 b. (To) da je Marija otišla, ja [znam / sumnjam / već shvatih]. (BCMS)
 it that has M. left I regret doubt already realized
 'That Marija left, I [know/doubt/deny/just realized].'

The grammaticality contrasts in (24)-(25) thus further confirm the analysis according to which volunteer stance verbs constitute CP structures, while others project a DP on top of CP.

The present analysis is also supported by data pertaining to syntactic islandhood. Simple CP structures are not expected to bring about syntactic island effects, so we can predict that volunteer-stance verbs should freely allow for wh-extractions. In the examples below, we can see that this is indeed the case:

(26) a. Kakvo_i misliš, će e kupil t_i? (Bg)
 what you-think that has bought
 'What do you think that he bought?'
 b. Kogai_i kaza, će e došla t_i?
 when you-said that has come
 'When did you say that she came?'

(27) a. Što_i si rekao da ste kupili t_i? (BCMS)
 what you-have said that you-have bought
 'What did you say that you have bought?'
 b. Kad_i misliš da je otišla t_i?
 when you-think that has left
 'When do you think that she left?'

¹⁴ Factive-emotive verbs like *sužaljavam* 'regret' are not included in this diagnostic because the possible presence of a PP projection on top of DP (see footnote 13 above) complicates the checking of the D feature in T, thus rendering the fronting of factive-emotive complements to the subject position more difficult (regardless of the fact that they contain a DP projection on top of the clausal CP).

On the other hand, DP structures are expected to exhibit strong island effects, given the more generalized ban on wh-extractions from (definite) DP constituents. This prediction is confirmed in the case of non-factive referential verbs and true factives in Bg and BCMS.

(28) a. * Kakvo_i [sūžaljavaš / otričaš], če ste prodali t_i? (Bg)
 what you-regret you-deny that you-have sold
 ‘*What do you [regret/deny] that you sold?’

 b. * Kūde_i [sūžaljavaš / otričaš], če ste otišli na počivka t_i?
 where you-regret you-deny that you-have gone on vacation
 ‘*Where do you [regret/deny] that you went on vacation?’

(29) a. * Što_i [žališ / poričeš] da ste prodali t_i? (BCMS)
 what you-regret you-deny that you-have sold

 b. * Gdje_i [žališ / poričeš] da ste otišli na odmor t_i?
 where you-regret you-deny that you-have gone on vacation

As for semi-factive predicates, they exhibit mixed properties in relation to islandhood as well. They bring about strong island effects when they are headed by the overt pronominal *tova/to* and no island effects when they are not headed by these pronominals, as shown in the South Slavic examples below.¹⁵

(30) a. Kakvo_i [otkri / razbra] (*tova), če e prodala t_i? (Bg)
 what you-found out you-realized it that has sold
 ‘What did you [find out/realize] that she sold?’

 b. ? Kūde_i [otkri / razbra] (*tova), če se provežda koncertūt t_i?
 where you-found out you-realized it that is-held concert
 ‘Where did you [find out/realize] that the concert is being held?’

(31) a. Što_i si [otkrio / shvatio] (*to) da je prodala t_i? (BCMS)
 what you-have found out realized it that has sold

 b. ? Gdje_i si [otkrio / doznao] da se održava koncert t_i?
 where you-have found out learned that is-held concert

I will argue that the contrasts in terms of island effects observed with semi-factive verbs in (30)-(31) are due to their ambiguous syntactic status. When complements to these verbs are headed by the overt *tova/to* item, they project a DP projection on top of CP and hence constitute a strong island domain, just like true factives or non-factive referential verbs. In the absence of *tova/to*, however, semi-factives do not project a DP on top of CP and hence do not bring about island effects, patterning in this sense with CP complements to volunteer-stance verbs in (26)-(27).

The analysis proposed above can be further strengthened if we return once again to the data pertaining to referentiality discussed in Section 3. The claim that I will make in this context is that only CP complements can be used to bring forth new information, whereas DP complements are necessarily grounded in discourse and can thus only be used to reiterate old information. Hence, CP-selecting volunteer-stance verbs can freely appear in answers to

¹⁵ Wh-adjunct extractions are not as freely acceptable with complements to semi-factive verbs as they are with volunteer-stance verbs (hence the question mark in front of the examples in [30b]-[31b]), possibly due to pragmatic constraints. Nevertheless, there is a clear contrast between semi-factive complements headed by *tova/to* and those that are not headed by these items in relation to syntactic islandhood. Only in the former case do we observe clear strong island effects.

questions (32a-b), whereas DP-selecting non-factive referential verbs and true factives cannot (32a'-b') (the relevant examples from Section 3 are repeated below).

(32) When does the game begin?

- a. [Mislja / vjarvjam], če započva v 7 časa. (Bg)
I-think I-believe that begins at 7 o'clock
- b. [Mislim/vjerujem] da počinje u 7 sati. (BCMS)
I-think I-believe that begins at 7 o'clock
- a.' # [Sūžaljavam/otričam], če započva v 7 časa. (Bg)
I-regret I-denry that begins at 7 o'clock
- b.' # [Žalim / poričem] da počinje u 7 sati. (BCMS)
I-regret I-denry that begins at 7 o'clock

Semi-factives, on the other hand, exhibit mixed behavior in this context as well: they can bring forth new information in an answer to a question when they introduce a CP structure, but this is no longer possible when they introduce a DP structure, i.e. when they are headed by the item *tova/to*.

(33) When does the game begin?

- a. [Osūzna / otkrih] (*tova) da započva v 7 časa.
I-learned I-found out it that begins at 7 o'clock
- b. [Saznah / otkrih] (*to) da počinje u 7 sati.
I-learned I-found out it that begins at 7 o'clock
'I [learned/found out] that it begins at 7 o'clock.'

The pronominal items *tova* and *to* can thus be seen as overt markers of referentiality, since all complements that introduce these elements on top of the clausal CP are necessarily referential and cannot bring forth new information.

At this point, I can propose the relevant structural descriptions for all the verb groups previously listed in (i)-(iv), which can explain their contrasting properties in relation to (non)referentiality.

(34) **Syntax of referentiality**

- (i) Volunteer-stance verbs (e.g. *say*, *claim*, *think*, *believe*, etc.).
V [CP [TP [vP]]] → **non-referential**
- (ii) Non-factive referential verbs (e.g. *doubt*, *deny*, *disprove*, *debunk*, etc.)
V [DP DREF [CP [TP [vP]]]] → **referential**
- (iii) Semi-factive verbs (e.g. *realize*, *find out*, *discover*, *learn*, *prove*, etc.)
 - a. V [CP [TP [vP]]] → **non-referential**
 - b. V [DP DREF [CP [TP [vP]]]] → **referential**
- (iv) True factive verbs (e.g. *regret*, *lament*, *be glad*, *be happy*, *be surprised*, etc.)
V [DP DREF [CP [TP [vP]]]] → **referential**

The referential properties of a given clause are encoded on the **DREF** head in (34), which grounds the proposition within the previous discourse. This analysis can, therefore, explain why complements that project a DP projection on top of CP are referential, whereas those that do not project a DP are non-referential: the **DREF** head is present in the former case and absent in the latter. Nevertheless, while the syntactic analysis in (34) can account for the observed contrasts between different predicates and their complements in relation to (non)referentiality,

it is not yet able to account for the contrasts that they were shown to exhibit in relation to (non)factivity. The latter will be explained in the remaining parts of the paper.

I will argue that the only group of predicates that can encode factivity at the level of syntactic structure are true factives (i.e. factive-emotive verbs) in (iv). The factive-type readings associated with other groups of predicates are a matter of the verb's lexical semantics (this being the case especially with semi-factive verbs) or are determined contextually, which also means that factive/presuppositional readings in such cases are more easily cancelled in different types of syntactic environments (as we already observed previously). The rest of the paper will focus on the syntactic realization of factivity in complements to true factive predicates.

I will argue that factive readings are encoded at the level of CP, via a dedicated CFACT head, which was already postulated by Kratzer (2006). The C-head in question can have different morphological manifestations across languages. For instance, Ihsane & Shim (2017) have demonstrated that the C-head embedded under factive emotive verbs like *regret* in English differs from the C-head selected by other predicates (including semi-factives) because the former, unlike the latter, disallows complementizer drop (COMP-drop). Note the grammaticality contrast in (35):

(35) a. I [think/believe/realize/know] (that) she left.
 b. I regret *(that) she left.

A similar contrast with regard to COMP-drop was noted in Italian as well, specifically in relation to subjunctive complements (Giorgi 2009). The deletion of the Italian complementizer *che* is possible if the subjunctive is selected by propositional-attitude verbs (e.g. *pensare* 'think', *credere* 'believe', etc.) but not if it is selected by factive-emotive verbs (e.g. *dispiacere* 'dislike', *rimpiangere* 'regret', etc.), as shown below:

(36) a. [Penso / credo] (che) sia venuto. (It)
 I-think I-believe that has.SUBJ arrived
 'I [think/believe] (that) he has arrived.'
 b. [Mi dispiace / rimpiango] *?(che) sia partito.
 I-dislike I-regret that has.SUBJ left
 'I [dislike/regret] that he left.'

The English and Italian data in (35)-(36) would thus seem to suggest that factive-emotive predicates constitute a special syntactic verb class in these languages, selecting complementizers that are resistant to COMP-drop.

I will argue that the same is true of factive-emotive predicates (or true factives) in South Slavic as well. The special status of these predicates in languages like Bg and BCMS is not manifested in relation to COMP-drop (which is not possible in South Slavic in general), but through the selection of distinctive factive complementizers. The languages under study feature dedicated factive complementizers, namely the Bg *deto* and the BCMS *što*, which are (almost) exclusively selected by factive-emotive verbs.¹⁶ Note the grammaticality contrasts between true factives (37) and semi-factives (38) below:

¹⁶ There are a few rare exceptions (noted by Baunaz [2015, 2016], among others) where the factive complementizer is occasionally selected by some verbs that are not true factives, this being the case, for instance, with the Bg verb *pomnjam* 'remember'. These exceptions (which are not found in BCMS and are very rarely found in Bg) will not be dealt with in the present paper, but we could speculate that the factive *deto* in such cases is not selected by the matrix predicate per se (given the present assumption that only true factives can select this item in Bg) but is introduced into the structure through some other type of syntactic mechanism. A more precise analysis in this context will be left for future work.

(37) a. [Sūžaljavam/iznenadan sūm], [deto / če] e zaminala. (Bg)
 I-regret surprised I-am that.FACT that has left

b. [Žalim / iznenaden sam] [što / da] je otišla. (BCMS)
 I-regret surprised I-am that.FACT that has left
 'I [regret/am surprised] that she left.'

(38) a. [Znam / razbrah / otkrih] [*deto / če] e zaminala. (Bg)
 I-know I-realized I-found out that.FACT that has left

b. [Znam / shvatih / otkrih] [*što / da] je otišla. (BCMS)
 I-know I-realized I-found out that.FACT that has left
 'I [know/realized/found out] that she left.'

As we can see in the examples above, true factive verbs can optionally select either a dedicated C_{FACT} complementizer (i.e. Bg *deto* and BCMS *što*) or a default declarative complementizer (i.e. Bg *če* and BCMS *da*), whereas semi-factives (as well as all other verb groups that have been analyzed here) can only select the latter.

Nevertheless, the fact that factive-emotive verbs are underspecified with regard to the type of complementizer (factive vs. non-factive) they select indicates that even these predicates do not inherently encode factivity at the level of their lexical semantics.¹⁷ The true marker of factivity is the C_{FACT} complementizer, which introduces truth-presuppositional readings that are uncancelable regardless of context. The decisive contribution of the complementizers *deto* and *što* to factive readings can be observed, for instance, if we look at certain emotive-type verbs which are ambiguous in the sense that they allow both factive and non-factive interpretations (e.g. Bg *oplakvam se* 'complain' or BCMS *brinuti se* 'to worry'). When these predicates select the default declarative complementizer, their presuppositional reading can be cancelled by the speaker (39a)/(40a), but when they select the factive complementizer, this is no longer possible (39b)/(40b).¹⁸

(39) a. Bašta mi se oplakva, če zakasnjavam večer (no toj greši, az vinagi father my complains that I-am-late evening but he is-wrong I always se pribiram predi 10 časa). (Bg)
 return before 10 o'clock

b. Bašta mi se oplakva, deto zakasnjavam večer (# no toj greši, az vinagi father my complains that.FACT I-am-late evening but he is-wrong I always se pribiram predi 10 časa).
 return before 10 o'clock
 'My father complains that I return late in the evenings (but I always come back before 10 o'clock).'

(40) a. Ivan se brine da je Marija bolesna (ali s njom je sve u redu). (BCMS)
 I. worries that is M. sick but with her is everything in order

b. Ivan se brine što je Marija bolesna (# ali s njom je sve u redu).
 I. worries that.FACT is M. sick but with her is everything in order
 'Ivan worries that Marija is sick (but she is completely fine).'

¹⁷ The reason why I use the label 'true factives' in relation to these predicates is not because they are inherently factive per se (see, for instance, [41]-[42]) but because they are the only group of predicates that can select the C_{FACT} complementizer, which is the true marker of factivity.

¹⁸ I am grateful to Iliyana Krapova for providing me the Bg example in (39).

Given that the predicates in (39)-(40) are underspecified with regards to factivity, the only true trigger of factive/presuppositional readings in this context is the factive complementizer, i.e. Bg *deto* and BCMS *što*.

Note, moreover, that even with factive emotive verbs such as *regret*, one can come up with certain contexts (admittedly somewhat contrived) where the factive/presuppositional reading associated with such predicates is cancelled out. One such instance is the context of ‘false belief’, where the speaker signals their disagreement with the matrix subject about the truthfulness of the proposition embedded under the factive-emotive verb, thus cancelling the presuppositional reading in this context (Klein 1975, Gazdar 1979, Baunaz & Puskás 2022, etc.). Nevertheless, this is once again only possible if the verb in question selects the default complementizer, whereas in the presence of the factive complementizer the presupposition is uncancellable regardless of context.¹⁹

(41) Edip pogrešno misleše, če e nanesäl smärtonosna rana i sǎžaljavaše,
Oedipus wrongly thought that has inflicted fatal wound and he-regretted
[če / # *deto*] e ubil strannika po pǎrja za Tiva. (Bg)
that that.FACT has killed stranger on road for Thebes
‘Oedipus wrongly believed that he inflicted a fatal wound and he regretted killing the
stranger on the road to Thebes.’

(42) Ivan pogrešno vjeruje da se Marija udala, i žalostan je [da / # *što*] više
I. wrongly believes that has M. married and sad is that that.FACT anymore
nije dostupna. (BCMS)
NEG-is available
‘Ivan wrongly believes that Marija got married and he is sad that she is no longer
available.’

Thus, once again, the true trigger of factivity and presuppositionality is the CFACT complementizer (*deto/što* in the case of South Slavic), not the factive verb per se.

I will conclude the analysis presented in this section with a brief foray into formal, possible-worlds semantics, which will allow me to provide a more precise account of the functioning of CFACT complementizers in relation to factivity. First of all, factive presuppositions differ from standard propositions when it comes to the type of world-anchoring they involve. Standard propositions are usually semantically analyzed as being anchored to a set of possible worlds, which contains both the worlds in which the proposition is true (let’s call them *p*-worlds) as well as worlds in which it is false (non-*p* worlds).²⁰ As a result, propositions can be judged as either true or false. In the context of embedded clausal complements, we can assume that the default declarative complementizer (e.g. Bg *če* or BCMS *da*) trivially anchors the proposition to this set of possible worlds, and then it is the lexical semantics of the matrix predicate and the overall context which determine whether the proposition is more likely to be true or false. Thus, for instance, with semi-factive predicates, there are more *p*-worlds than non-*p* worlds (and the proposition is thus more likely to be true), and this is even more so the case of true-factive verbs. Nevertheless, whenever these verbs select a default complementizer, there will always be a subset of non-*p* worlds allowing for the possibility to cancel the truth reading of the embedded proposition (even with true factive verbs, as we saw in [41]-[42] above).

However, when a complementizer is introduced under CFACT, we get a completely different type of world-anchoring. To be more precise, CFACT can be described as a ‘uniqueness

¹⁹ The example in (41) is adapted from Klein (1975) and the one in (42) is adapted from Baunaz & Puskás (2022).

²⁰ See Portner (2009) and the references therein.

operator', which reduces the set of possible worlds associated with the embedded proposition to a singleton set (in the sense of Schueler 2016) containing a single *p*-world, where the truth of the proposition necessarily obtains. This can explain why the truth-presuppositional reading associated with factive complements embedded under complementizers such as the Bg *deto* or the BCMS *što* can never be cancelled regardless of context. The world-semantics analysis briefly presented here still needs to be hashed out in more detail, but it provides a potentially promising avenue for further research into the semantics of factivity.

6. CONCLUSION

The main goal of the paper was to classify different types of predicates in relation to factivity and referentiality, and then develop a syntactic analysis that can account for the contrasting patterns that the various verb classes have been shown to exhibit in this context. The contrasts pertaining to (non)referentiality were explained by postulating a nominal D_{REF} head on top of the clausal CP whose function is to anchor the utterance to the previous discourse. The presence of this D-head in complements to referential verbs was demonstrated through a range of independent syntactic evidence (e.g. island effects or complement fronting to the matrix subject position, among others). Factivity, on the other hand, was shown to be encoded at the level of C via a dedicated C_{FACT} complementizer (realized as *deto* in Bg and *što* in BCMS), which functions as a uniqueness operator anchoring a given utterance to a single world where its truth necessarily obtains. The goal for future work will be to develop a more fine-grained syntactic and semantic analysis pertaining to the notions that were discussed here, as well as apply it to a broader range of languages.

REFERENCES

Abbott, B. (2000) "Presuppositions as Non-Assertions", *Journal of Pragmatics* 32, 1419–1437.

Abrusán, M. (2011) "Predicting the Presuppositions of Soft Triggers", *Linguistics and Philosophy* 34, 491–535.

Anand, P., J. Grimshaw & V. Hacquard (2017) "Sentence Embedding Predicates, Factivity and Subjects", in C. Condoravdi (ed.) *Lauri Karttunen Festschrift*, CSLI publications, Stanford, 1–29.

Baunaz, L. (2015) "On the Various Sizes of Complementizers", *Probus* 27(2), 193–236.

Baunaz, L. (2016) "Deconstructing Complementizers in Serbo-Croatian, Modern Greek and Bulgarian", in C. Hammerly & B. Prickett (eds.) *Proceedings of NELS* 46(1), Graduate Linguistics Student Association, Amherst, Mass., 69–77.

Baunaz, L. & G. Puskás (2022) *A Cross-Linguistic Approach to the Syntax of Subjunctive Mood*, Studies in Natural Language and Linguistic Theory, Springer.

Cattell, R. (1978) "On the source of Interrogative Adverbs", *Language* 54, 61–77.

de Cuba, C. (2007) *On (Non)Factivity, Clausal Complementation, and the CP-Field*, PhD dissertation, Stony Brook University.

de Cuba, C. (2017) "Noun Complement Clauses as Referential Modifiers", *Glossa* 2.1(3), 1–46.

de Cuba, C. & B. Ürögdi (2010) "Clearing up 'Facts' on Complementation", in J. S. Stevens (ed.) *University of Pennsylvania Working Papers in Linguistics* 16(1), 41–50.

Djärv, K. (2019) *Factive and Assertive Attitude Reports*, PhD dissertation, UPenn.

Djärv, K., C. Heycock & H. Rohde (2017) "Assertion and Factivity: Towards Explaining Restrictions on Embedded V2 in Scandinavian", in L. Bailey & M. Sheehan (eds.) *Order and Structure in Syntax I: Word Order and Syntactic Structure*, Language Science Press, Berlin, 3-28.

Egré, P. (2008) "Question-Embedding and Factivity", *Grazer philosophische Studien* 77(1), 85–125.

Gazdar, G. (1979) *Pragmatics: Implicature, Presupposition, and Logical Form*, Academic Press, New York.

Giorgi, A. (2009) "Toward a Syntax of the Subjunctive Mood", *Lingua*, 119, 1837–1858.

Haegeman, L. (2006) "Conditionals, Factivities and Left Periphery", *Lingua* 116, 1651–1669.

Heim, I. (1992) "Presupposition Projection and the Semantics of Attitude Verbs", *Journal of Semantics* 9, 183– 221.

Ihsane, T. & J. Y. Shim (2017) "A New Outlook on Complementizers", *Languages* 2/17, 1–18.

Karttunen, L. (1971) "Some Observations on Factivity", *Papers in Linguistics* 4(1), 55–69.

Karttunen, L. (1973) "Presuppositions of Compound Sentences", *Linguistic Inquiry* 4, 169– 193.

Karttunen, L. (1974) "Presupposition and Linguistic Context", *Theoretical Linguistics* 1(1-3), 181–194.

Kastner, I. (2015) "Factivity Mirrors Interpretation: The Selectional Requirements of Presuppositional Verbs", *Lingua* 164, 156–188.

Kiparsky, P. & C. Kiparsky (1970) "Fact", in D. Steinberg & L. Jakobovits (eds.) *Semantics*, Cambridge University Press, Cambridge, Mass., 345–369.

Klein, E. 1975. "Two Sorts of Factive üredicate", *Pragmatics Microfiche it*, 1.1, frames B5– C14.

Krapova, I. (2010) "Bulgarian Relative and Factive Clauses with the Invariant Complementizer *deto*", *Lingua* 12, 1240–1272.

Kratzer, A. (2006) "Decomposing Attitude Verbs", Talk given in honor of Anita Mittwoch, The Hebrew University of Jerusalem, Jerusalem.

Langendoen, T. D. & H. B. Savin (1971) "The Projection Problem for Presuppositions", in C. J. Fillmore & T. D. Langendoen (eds.) *Studies in Linguistic Semantics*, Holt, Rinehart and Winston, New York, 373–388.

Portner, P. (2009) *Modality*, Oxford University Press, Oxford.

Schueler, D. (2016) "Factivity and Complement-Types", *Studia Linguistica* 70(3), 297–335.

Simons, M. (2007) "Observations on Embedding Verbs, Evidentiality, and Presupposition", *Lingua* 117, 1034–1056.

Stalnaker, R. C. (1974) "Pragmatic Presupposition", in M. Munitz & P. Unger (eds.) *Semantics and Philosophy*, New York University Press, New York, 197–213.

Stalnaker, R. C. (2002) "Common Ground", *Linguistics and Philosophy* 25, 701–721.