

## Chapter 7

# The nature(s) of syntactic variation: Evidence from the Serbian/Croatian dialect continuum

Peđa Kovačević

University of Novi Sad

Tanja Milićev

University of Novi Sad

The paper reports on a study of the variation inside the Serbo-Croatian dialect continuum with respect to clitic placement, complements of modal verbs (infinitive/DA+present) and the use of *trebati* ‘need’ as either an experiencer verb or a simple transitive. Region and ethnicity accounted for a large portion of variation in the use of infinitives and DA+present with many speakers using these structures interchangeably. Next, we found that clitics are almost uniformly placed after the first phrase. The variation in the use of lexical *trebati* was confined to the Croatian portion of the sample. Our findings suggest that (i) infinitives and DA+present after modal verbs should be treated as roughly the same syntactic structure; (ii) variation in clitic placement should not be analyzed as an instance of sociolinguistic variation and deeper (linguistic) causes of variation should be pursued; (iii) *trebati* as a transitive verb appears in the Croatian variety only.

**Keywords:** syntactic variation, Serbo-Croatian dialect continuum, clitic placement, non-finite complements

## 1 Introduction

### 1.1 Syntactic variation: Theoretical framework

Recent theoretical approaches to syntactic variation have enabled us to form a more fluid picture of syntax (Adger 2006; Adger & Trousdale 2007; Adger &

Smith 2005). Updates on the rather rigid classical Principles and Parameters theory (Chomsky & Lasnik 1995) like Kayne's (2000) Microparametric approach or Kroch's (1994) Competing Grammars have struggled with the fact that syntactic variation can be quite free and apparently even optional in some cases. More recently, in line with more general theoretical advances, it has been argued that syntactic variation, with all its apparent fluidity, can be captured within the Minimalist Framework (Adger 2006; Adger & Trousdale 2007; Adger & Smith 2005).

This approach provides us with a way of looking at variation which predicts much greater freedom on the part of speakers to move between two different structures depending on the context. Furthermore, the approach is freed of the assumption that some speakers constantly move from one grammar to another as they produce different constructions.

Instead of relying on parameters and/or microparameters as explanatory mechanisms, Adger and colleagues assume that syntax is simply a set of uniform core operations applied to lexical items. What appears as syntactic variation, thus, arises when (i) there are two or more ways of pronouncing the same structure or (ii) when there are different uninterpretable morphosyntactic features on competing lexical items. Adger (2006) illustrates this with an example of different T heads that can be found in Standard English and dialects like Buckie English and others, which give rise to different spellouts of the auxiliary *be*. While in Standard English, T is sensitive to agreement and spells out the agreement patterns morphophonologically, in non-standard dialects, *be* is either completely insensitive to agreement (i.e. bears no uninterpretable phi-features) or simply does not spell out reflexes of agreement in the same way as in Standard English. Either way, a speaker can have both lexical items (T heads) in their mental lexicon and depending on which one they choose, the output will vary. The way the speaker employs these different lexical items is determined by sociolinguistic factors in the sense of Labov (1972).

This approach provides us with a way of looking at variation, which predicts much greater freedom on the part of speakers to move between two different structures depending on the context. Furthermore, the approach is freed of the assumption that some speakers constantly move from one grammar to another as they produce different constructions.

The Serbo-Croatian dialect continuum provides very useful testing ground for theories of syntactic variation. Our primary goal in this paper is to present some data from an empirical study of three instances of syntactic variation in this dialect continuum in order to arrive at a clearer factual description of the phenomena at hand. We will also provide sketches of formal analyses of these three structures, which will show that the approach developed by Adger and his co-workers is a very useful theoretical tool when it comes to explaining the observed data.

## 1.2 Syntactic variation in the Serbian/Croatian dialect continuum

It is a well-known fact that the differences between Serbian and Croatian standard varieties belong mostly to the lexicon and the domain of pronunciation (Corbett & Browne 2009; Bailyn 2010, *inter alia*). The most prominent differences in the realm of pronunciation have to do with the way in which speakers of these two varieties pronounce words that used to contain the so-called yat sound in the older varieties of the language. In modern Serbian, this sound is pronounced as /e/, while in modern Croatian it is either /je/ or /ije/, depending on the length of the earlier vowel. Based on these different pronunciations, the two standard varieties are also called Ekavian and Ijekavian. In terms of differences in vocabulary items, one can mention that due to historical factors the Croatian variety tended to borrow more from German, Czech and other languages of Central Europe, while Serbian contains more borrowings from Turkish and other languages of the Balkans (Corbett & Browne 2009).

Syntactic variation in this dialect continuum seems limited to just a few potential cases. One of the best known points of difference has to do with the structure of non-finite verbal complements. Example (1a) illustrates the option of infinitives functioning as complements of modal verbs while in (1b), the modal is followed by the so-called DA+present structure. Standard grammars of Croatian draw a sharp distinction between the Serbian DA+present option and the Croatian infinitives (Katičić 1986). However, Bailyn (2010) provides some empirical evidence to the effect that both varieties allow both options and infinitives are simply more common in Croatian.

- (1) a. Ivan mora **pojesti** večeru.  
Ivan must eat.INF dinner  
'Ivan must eat his dinner.'
- b. Ivan mora **da pojede** večeru.  
Ivan must DA eat.PRES.3SG dinner  
'Ivan must eat his dinner.'

Another area of potential syntactic variation would be the positioning of clitics. When it comes to clitics, standard Croatian grammars prescribe placing the clitics after the first word, a rule that is sometimes referred to as the 2W rule (Katičić 1986; for criticism see Peti-Stantić 2009). This rule is illustrated in (2b). In Serbian, the most neutral rule is to place the clitics after the first phrase, a rule known as the 2P rule (2a). Corbett & Browne (2009) suggest that the 2W rule is less common in the context of the clitic-second phenomenon because under 2W, the clitic cluster splits a constituent.

- (2) a. Pravi igrač **je** došao.  
           true player AUX.CL come  
           ‘A true player has come.’  
       b. Pravi **je** igrač došao.  
           true AUX.CL player come  
           ‘A true player has come.’

Regarding the verb *trebati* ‘need’, we find that Standard Croatian grammars recognize its existence as a transitive verb (3a), while in Serbian, it appears only as an experiencer verb (3b). Also, as a modal verb, in Serbian *trebati* is prescribed as being always impersonal (4b), as opposed to Croatian (4a).

- (3) a. Ivan **treba** knjigu.  
           Ivan.NOM need.3SG book.ACC  
           ‘Ivan needs (some) cheese’  
       b. Ivanu **treba** knjiga.  
           Ivan.DAT need.IMP book.NOM  
           ‘Ivan needs (some) cheese.’  
       (4) a. Deca **trebaju** otići.  
               children.NOM.SG need.3PL go.INF  
               ‘Children need to go.’  
           b. Deca **treba** / \***trebaju** da odu.  
               children.NOM.SG need.IMP need.3PL DA go.3PL  
               ‘Children need to go.’

Standard grammars of both Croatian and Serbian often focus on essentially eliminating the variation and prescribing one option as “more natural” for a given variety. Therefore, in a sense, they present an overly rigid either–or, binary picture of variation in these domains.

In order to make sense of the variation in these domains, one needs to have a clear picture of the underlying facts, which we claim are not correctly represented in descriptive grammars. Therefore, the primary aim of this research is to provide some empirical insight into the nature of variation in these three aspects. Next, we will argue that the data point towards the view of variation proposed by Adger & Trousdale (2007) and Adger & Smith (2005). Finally, we will suggest ways of analyzing these constructions formally based on the implications that arise from this particular view of variation.

## 2 Empirical data: Production study

The enumerated instances of potential syntactic variation in the Serbo-Croatian dialect continuum were confirmed by a simple production study.<sup>1</sup> The production task consisted of a written survey that elicited structures like (1)–(3). The non-finite complements of modals were elicited by means of sentences like (5) where the target verb appears in the first part of a compound sentence in its finite form. In the second part of the sentence, the same verb is supposed to appear in its non-finite form after a modal, but a blank is given in its stead. The participants were instructed to fill in the blank with the form of the underlined verb that they found most suitable. They were also instructed not to leave out the verb because those sentences are grammatical even when the verb is elided. There were 20 sentences in total, and the targeted non-finite structures were placed in the contexts of modals *moći* ‘can’, *morati* ‘must’, the phasal verb *početi* ‘start’, and the verb *želeći* ‘want’. Five target sentences were dedicated to each of these contexts.

- (5) Milan je pojeo salatu, a Ivan još mora \_\_\_\_\_ desert.  
 Milan is ate salad, while Ivan still must \_\_\_\_\_ dessert  
 ‘Milan ate the salad while Ivan still has to (eat) the dessert.’

When it comes to the variations in clitic placement, the task was to shift sentences like (6a) into past tense (6b). As can be seen in the examples in (6), the sentence in the present tense does not contain clitics, but in the past tense, the auxiliary clitic *je* ‘is’ is necessary. However, the position of the clitic can be varied as indicated in the example. It can either come immediately after the demonstrative *ta* ‘that’, or it can come after the subject noun phrase *ta gospođa* ‘that lady’, in accordance with 2W or 2P rules respectively. There were 12 target sentences in total and the sentences were organized into four groups according to the type of the prenominal modifier (demonstrative, descriptive adjective, possessive or demonstrative adjective). Each modifier appeared in all three genders (masculine, feminine, neuter).

- (6) a. Ta gospođa pravi kolače.  
       that lady makes cookies  
       ‘That lady makes cookies.’  
       b. Ta {je} gospođa {je} pravila kolače.  
       that AUX.CL lady AUX.CL made cookies  
       ‘That lady made cookies.’

<sup>1</sup>A detailed description of the design, including all the experimental items, can be found at <https://osf.io/m5feh>.

Finally, the verb *trebati* ‘need’ was elicited by means of asking a question where the most likely response will be a sentence containing this verb. However, the crucial thing to worry about was avoiding the use of this verb in the question itself because the way the verb is used in the question would have a great impact on how it would be used in the answer. Because this was a written task, it was possible to provide one or two sentences in the way of context and then ask the question like (7a) using a verb other than *trebati*, but also state that the verb *trebati* should be used in the answer.

- (7) a. Šta je još potrebno Petru?  
what is else needed.IMP Peter.DAT  
‘What else does Peter need?’  
b. Petar treba olovku.  
Peter.NOM need.3SG pencil.ACC  
‘Peter needs a pencil.’  
c. Petru treba olovka.  
Peter.DAT need.IMP pencil.NOM  
‘Peter needs a pencil.’

The possible answers to the question in (7a) were either (7b) or (7c). The choice of one over the other would reveal the way the participant uses the verb *trebati* in his or her everyday speech. Because we studied the variation in the use of this one verb only, we had only three target sentences that we wanted to elicit.<sup>2</sup>

When it comes to the choice of participants, we were interested in the way in which geographic location and ethnicity influenced the use of these constructions. Our sample consisted of 120 participants from Serbia and Croatia, ages 16-19. They were divided into four groups with 30 participants each. One group consisted of 30 students attending the so-called gymnasium school (*gimnazija*) in Zagreb. One group was located in the town of Ruma, roughly 60 kilometers west of Belgrade. This group also consisted of 30 gymnasium students. Finally, there were two groups in the town Subotica, in the north of Serbia, on the border with Hungary. The reason why we had two groups in this town was because in Subotica, there was the option of varying the ethnicity of the participants while controlling for their geographic location. Namely, the gymnasium in this town

---

<sup>2</sup>An anonymous reviewer points out that the presence of a dative argument in the elicitation question could have primed the subjects to also use a dative in the response with the verb *trebati*. This might have reduced the number of transitive uses. The fact that the Zagreb group still largely opted for the transitive *trebati* (as opposed to groups from Serbia) shows that the possible priming effect was not nearly strong enough to suppress the transitive use.

has a Croatian track alongside the Serbian track. What this means is that students have the option to enroll in classes that are taught in standard Croatian, and a number of ethnic Croats choose this option. In order to vary the ethnicity of the students while controlling for geographic location, we created one group of 30 students from the Croatian track and one group of 30 students from the Serbian track.

It is important to note that our sample was constructed in such a way as to compare the dialect spoken in Zagreb with the dialects spoken in Vojvodina, the Northern Province of Serbia. Zagreb was taken as a benchmark representing a dialect close to the Croatian standard (the participants were students of the Classical Gymnasium in Zagreb, a very prestigious school with a strong focus on languages). Towns in Vojvodina, on the other hand, were of interest to us because they represent the kind of gray area between the Croatian and the Serbian standard where one can zoom in on the speakers who speak neither of the standards but are quite close to both of them at the same time.

Finally, it should be pointed out that this was a pilot study into the vast realm of syntactic variation. We believe, though, that it gives a good starting point towards the understanding of patterns in variation when it comes to the syntactic structures we focused on.

### 3 Findings

The empirical data that we obtained pointed to quite different patterns of variation in the three structures under investigation. Concerning the variation in non-finite complements, we compared our groups based on the number of infinitives that each participant produced. In Figure 1, mean values for the number of infinitives are given for each group. The results from the groups from Subotica are

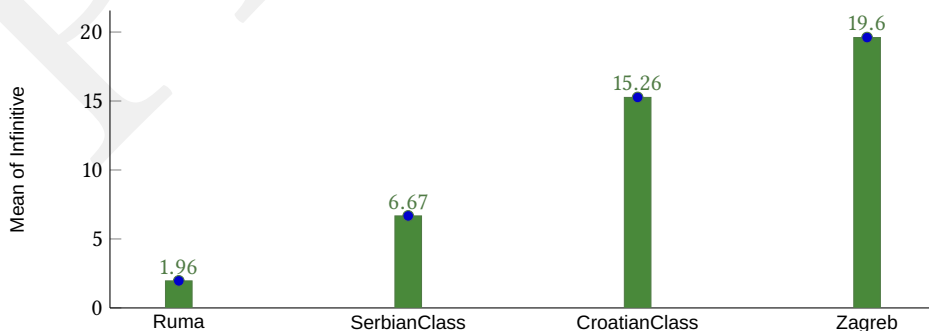


Figure 1: The average number of infinitives across groups

given in the middle of the graph with SerbianClass standing for the group made up of students attending the Serbian track and CroatianClass stands for students attending the Croatian track.

As the graph in Figure 1 suggests, there are important differences in the use of infinitives as non-finite complements across these groups. These differences are statistically significant (LR  $p < 0.01$ ,  $r^2 = 0.63$ ). Despite the fact that the group in Zagreb used infinitives almost exclusively, we can conclude that these structures can vary quite freely in the production of a significant number of speakers. The histogram in Figure 2, which shows how the use of infinitives was distributed across the entire sample provides a deeper insight into the nature of the variation in this area.

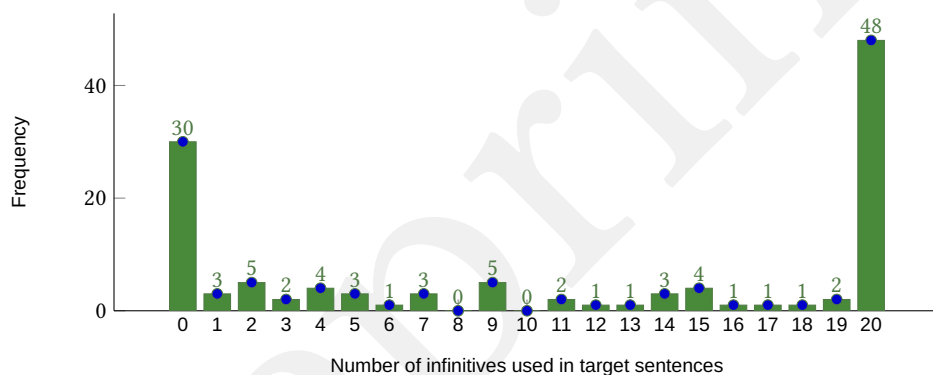


Figure 2: Frequency distribution for the number of infinitives in the entire sample

In Figure 2, the x-axis represents the numbers of infinitives used in the target sentences while the height of the bar shows how many participants who produced a particular number of infinitives there were. As the graph shows, a large portion of the participants either used infinitives throughout, or they systematically avoided them. The height of the bar above zero on x-axis shows the portion of participants who did not use infinitives at all while the height of the bar above 20 on the same axis indicates the share of the subjects who used infinitives only. However, there is also a sizable portion of the sample where these structures are in quite free variation. In other words, for many participants there were no clear preferences for either infinitive or DA+present. A closer look at the surveys done by some of these participants reveals no discernible pattern or context-dependent preference for one of the structures.

When it comes to the variation in clitic placement, we obtained very different



results. In our survey, there were 12 target sentences eliciting one or the other clitic placement option. In Figure 3, we plotted the mean numbers of sentences in which 2P rule was observed. The means are given for each of the four groups.

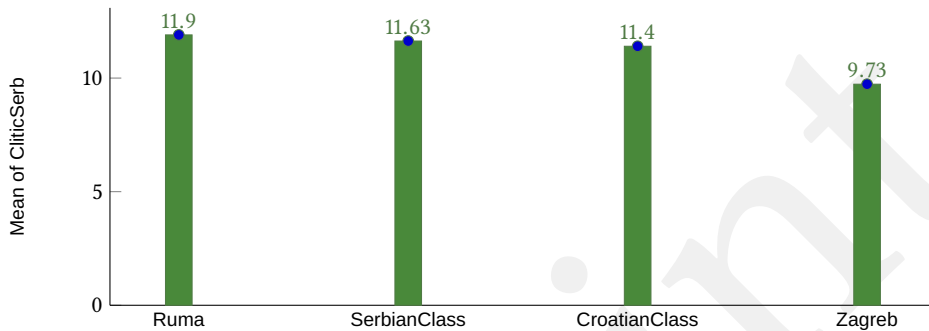


Figure 3: The average number of applications of 2P rule across groups

Simply by inspecting the graph visually, one can notice that the pattern of variation was different from what was observed with non-finite complements. The mean values for each of the groups are close to the maximum of 12, and Linear Regression found no statistical difference among the groups ( $p = 0.205$ ,  $r^2 = 0.0136$ ). A more detailed look at the surveys reveals that a very small number of participants did produce several instances of the 2W rule, but these were rather marginal as the preponderance of participants in all four groups used the 2P rule only.

Turning now to the variation in the use of the verb *trebati*, we can say that whatever variation there is in the use of this verb, it is confined to the Croatian variety. All the participants from Serbia (both groups from Subotica and the group from Ruma), used this verb in its experiencer-like form. There were no instances of this verb used as a simple transitive in these three groups. On the other hand, we found that there is substantial variation in the use of this verb within the group from Zagreb. About a third of the elicited utterances containing the verb *trebati* were characterized by the simple transitive use (the mean value was 1.06 with 3 being the maximum). Curiously, it was not the case that out of 30 participants approximately a third used *trebati* as a transitive verb exclusively and the remaining 20 participants used this verb only in its experiencer version. The instances of *trebati* as a transitive verb were much more distributed within the group with some speakers using this verb two times as an experiencer verb and once as a transitive one. Of course, there were also those who produced two sentences with a transitive *trebati* and one with its experiencer-like counterpart.

Crucially, the outcome was that, in fact, only less than a third of the participants from Zagreb consistently used *trebati* as an experienter verb with no instances of its transitive version.

## 4 Analysis

Armed with these empirical insights about the patterns of variation with these three constructions, we can turn to the question of what these insights can tell us about their underlying structure. Also, we might be able to derive some suggestions as to the broader theoretical questions dealing with the nature of syntactic variation hinted at in the introduction. These will be the topics of this section.

### 4.1 Infinitive vs. DA+present

On a general note, one can say that the DA+present construction has received much more attention in the syntactic literature than infinitives (Todorović 2012; Mišeska-Tomić 2004). The fact that these two structures can be found in virtually free variation is rarely addressed (see Belić 2005 for exceptions). Todorović & Wurmbrand (2015) note that the *da* particle found in DA+present constructions can function as a complementizer (8), a modality marker (9) and finiteness marker on *v* (10).

- (8) Jovan je            tvrdio da čita            knjigu.  
Jovan AUX.3SG claimed DA read.PRES.3SG book  
'Jovan claimed to be reading the book.'  
(from Todorović & Wurmbrand 2015)

- (9) Jovan je            odlučio da spava            u garaži.  
Jovan AUX.3SG decided DA sleep.PRES.3SG in garage  
'Jovan decided to sleep / that he would sleep in the garage.'  
(based on Todorović & Wurmbrand 2015)

- (10) Marko je            počeo da radi            zadatak.  
Marko AUX.3SG started DA do.PRES.3SG homework  
'Marko started doing his homework.'

Even though under certain conditions (9) would also allow an infinitive after the main verb, our study focused on structures like (10), where infinitive alternates

with DA+present most clearly.<sup>3</sup> Todorović & Wurmbrand (2015) treat infinitives as bare VPs based on the fact that they seem to be unable to assign accusative case, typically associated with the causative *v*. That way, they postulate a syntactic difference between these two structures because sentences like (10) contain a full *v*P at least. Under their account sentences like (11), where infinitive is used as the complement of the phasal verb, should have a bare VP in the embedded part. They claim that the accusative case on the object of the infinitive is assigned by the matrix verb.

- (11) Marko je počeo raditi zadatak.  
 Marko AUX.3SG started do.INF homework  
 ‘Marko started doing his homework.’

However, because the data illustrate the possibility of completely free variation, we will propose that both “low-DA”, corresponding to (10), and infinitives have the same structure: both are *v*Ps.<sup>4</sup> If infinitives and DA+present were truly different structures, one would not expect to find speakers who use them interchangeably, as we did. A deeper structural difference of the *v*P / VP kind would give rise to clear preferences for one structure over the other either across regional varieties or, at the very least, across individual speakers.

There also might be some syntactic evidence against the claim that infinitives are merely VPs. The main piece of evidence is the availability of accusative case with infinitives in contexts where it is difficult to argue that the accusative is assigned by the matrix verb: copular constructions (12) and impersonals (13). If the ability to assign accusative is taken as a diagnostic, we should conclude that infinitives, like the “low-DA” structures are *v*Ps.

- (12) Položiti matematiku je teško.  
 pass.INF math.ACC is difficult  
 ‘It is difficult to pass the math exam.’

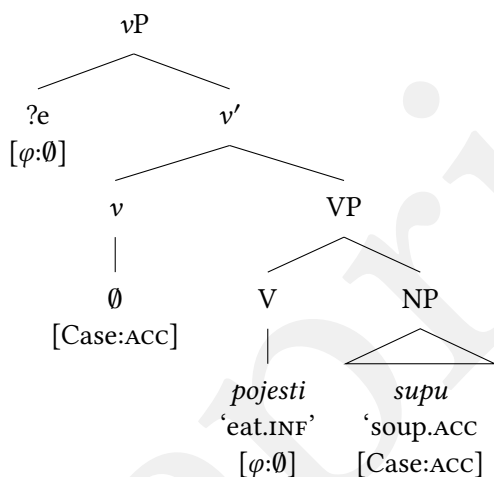
<sup>3</sup>Example (8) does not allow the alternation with infinitives, while the tense of the embedded clause can be varied, and the reference of the subject of the embedded clause is not tied to the reference of the matrix clause subject. Sentences like (9) allow infinitives and DA+present after the main verb only if the subjects of the matrix structure and the embedded structure are (referentially) the same. Having a (referentially) different subject is possible, but with DA+present only. Sentences like (10) never allow referentially different subjects in the embedded and the matrix part. We leave the variation of infinitives and DA+present in sentences like (9) for further research.

<sup>4</sup>By using the term “free variation”, we refer to structural alternations that have no consequences for the semantics and pragmatics of the sentence as a whole.

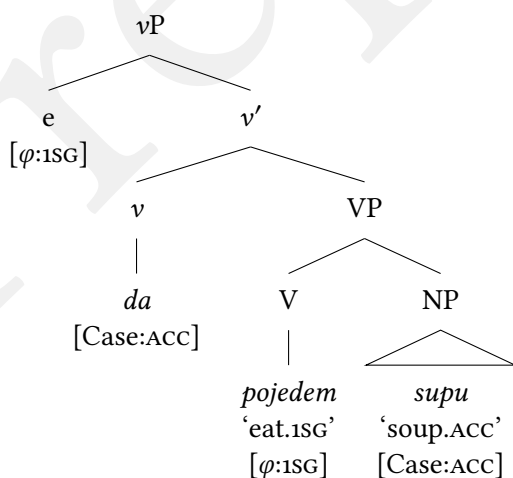
- (13) Trebalo je pojesti supu.  
 needed.3SG.N AUX.3SG eat.INF soup.ACC'  
 'One was supposed to eat soup.'

Once both infinitives and DA+present are reduced to essentially the same structure (i.e.  $vP$ ), we can look at them as simply different instantiations of the same  $v^0$ . The proposed structures for infinitives and DA+present are in (14).

- (14) a. **Infinitive**



- b. **DA+present**



If infinitives and DA+present are merely different instances of the same  $v^0$ , we can expect the kind of variation that we observed in our study. In line with the general view in Adger & Smith (2005), we can assume that there are some speakers whose mental lexicons contain both of these  $v$  heads, which is why they can use them interchangeably.

The structures in (14) raise one additional problem. Namely, it is unclear what the status of the embedded subject with infinitives and DA+present should be. Although both DA+present and infinitives are subject to the same constraints regarding the interpretation of the null subject (obligatory control, sloppy reading only, etc.), in impersonal constructions, we note a clear asymmetry with respect to the impersonal (reflexive) morpheme *se*. With DA+present, *se* (and in fact all kinds of pronominal clitics) obligatorily stays inside the  $vP$ , whereas with infinitives, it shows up with the matrix verb; see (15).

- (15) a. Moglo            je            da *se*    peva.  
           can.PAST.3SG.N AUX.3SG DA REFL sing.3SG  
           ‘It was possible to sing.’  
       b. \*Juće            *se*        moglo        da peva.  
           yesterday REFL can.PAST.3SG DA sing.3SG  
           Intended: ‘It was possible to sing yesterday.’  
       c. Moglo            *se*        pevati.  
           can.PAST.3SG.N REFL sing.INF  
           ‘It was possible to sing.’

While the behavior of *se* with DA+present supports our assumption for the existence of a null element in Spec $vP$  (which needs to be targeted/“switched off” by *se* impersonalization), the fact that *se* surpasses the infinitive predicate poses a problem for the uniform structural treatment of the infinitive and DA+present, and brings into the question the postulation of the  $vP$  layer in infinitives.<sup>5</sup> It is also possible that infinitives simply lack Spec $vP$ , which would still retain structural uniformity.

Wurmbrand (2003) and Todorović & Wurmbrand (2015) argue that there is no PRO with (restructuring) infinitives, i.e. that infinitives lack a syntactic subject altogether, and that interpretation comes from the matrix subject. However, a simpler way of capturing the relevant facts would be by postulating a difference in terms of the presence/absence of Spec $vP$  rather than saying that infinitives lack the  $vP$  layer completely. Again, saying that there is no  $vP$  with infinitives would leave sentences like (12) and (13) unexplained.

<sup>5</sup>Krapova’s (1999) analysis of a structure virtually identical to DA+present also assumes the existence of PRO in those contexts.

Impersonal contexts again provide us with evidence that the interpretation of the external argument of the infinitive is dependent on interpretation of the matrix predicate subject. Namely, infinitives are only possible with impersonal *se*. If *se* is absent, as in (16), only predicates without a referential subject (e.g. weather verbs, such as *grmeti* ‘thunder’ in (16c)) are possible.<sup>6</sup> No such restrictions hold for DA+present (16b), where *se* obviously takes care of getting the proper interpretation for the embedded predicate (indefinite, human).

- (16) a. \*Moralo / Moglo je pevati.  
must.PAST.N can.PAST.N AUX.3SG sing.INF  
Intended: ‘One had to / could sing.’
- b. Moralo / Moglo je da se peva.  
must.PAST.N can.PAST.N AUX.3SG DA SE sing.3SG  
‘One had to / could sing.’
- c. Moralo / Moglo je grmeti.  
must.PAST.N can.PAST.N AUX.3SG thunder  
‘There must / could have been thunder.’

Curiously, copular constructions and impersonal *trebati* ‘need’, which also lack an overt matrix subject, show no restrictions with respect to the infinitive (cf. (12) and (13)). It is possible that in these contexts we are dealing with what Wurmbrand (2003) calls “non-restructuring” configurations. These configurations would then be *vP* infinitives, which are different from the restructuring (VP) ones found after modals and verbs such as *try* or *begin*. At this point, we cannot provide a definitive resolution of this issue. The crucial test for a true case of restructuring is the availability of long passive. However, long passive in Serbian is possible only with impersonal *SE*-passive (cf. Todorović & Wurmbrand 2015), while *BE*-passive is not allowed in these constructions. The examples in (17) show the unavailability of long *BE*-passive, see (17a) and (17b), with both infinitive and DA+present complements together with the acceptable *SE*-passive versions.

- (17) a. \*Ta pesma je započeta da se svira. BE-passive  
that song.F is started.PASS.PTCP.F DA REFL play.3SG  
Intended: ‘They started to play that song.’
- a’. Ta pesma se započela svirati. SE-passive  
that song.F REFL started.PAST.PTCP.F play.INF  
‘They started to play that song.’

<sup>6</sup>Presumably, the subject of these matrix predicates is a kind of expletive *pro*.

- b. \* Ta pesma je započeta                      svirati.                      BE-passive  
       that song.F is started.PASS.PTCP.F play.INF  
       Intended: ‘They started to play that song.’
- b’. Ta pesma je započela                      da se svira.                      SE-passive  
       that song.F is started.PAST.PTCP.F DA REFL play.3SG  
       ‘They started to play that song.’

We leave open the question of obligatory control and whether the “PRO interpretation” of the infinitive requires a syntactic position or not.

It should be noted that some speakers report subtle differences in meanings of these two constructions. Examples like (18) illustrate some of these subtle differences. For speakers from central Serbia, these examples can mean simply negated future. For many speakers from Vojvodina, however, these sentences mean the lack of volition with both present and future temporal reference. These speakers prefer to use the infinitive for the meaning of negated future.

- (18) Mi nećemo            to            da radimo.  
       we NEG.will.1PL that DA do.1PL  
       ‘We will not do that.’

In sum, once we assume that DA+present and infinitive are two versions of the same *v* head, it becomes possible to explain the variation between the two as a consequence of the roughly equal availability of these two heads in the mental lexicons of such speakers. Also, the reason why some speakers consistently use one and never the other would be because their mental lexicon contains only one variety. At this point, the suggestion is to treat them as the same underlying structure.

## 4.2 Clitics

Concerning the difference between the 2W and 2P rules in the placement of clitics, one can identify two approaches. In one of the views, these two options are the same in terms of their underlying form (Ronelle 2006; 2008; Yu 2008). The difference, then, stems from the application of two different phonological processes, one of which inserts the clitics after the first word while the other one inserts the clitics after the first phrase. Crucially, these phonological processes apply differently across the dialect continuum. The difference, thus, seems to be understood to be purely sociolinguistic. Anderson (2005), surprisingly, even suggests that the use of 2W rule is not possible in Serbian, where only the 2P rule can be found.

Other authors argue that this difference is not purely sociolinguistic in nature. For instance, Diesing et al. (2009) argue that sentences in which the 2W rule is used have a marked pitch contour, which suggests prosodic focus on the prenominal modifier. Moreover, such sentences are claimed to be felicitous only in contexts where the prenominal modifier is contrastively focused. The 2P rule, on the other hand, is applicable to broad focus contexts and is, thus, interpreted as unmarked. Bošković (2009) proposes different syntactic derivations for the two rules. In his view, the 2W rule is derived by left-branch extraction of the prenominal modifier which then functions as an anchor for the clitic. Our own intuitions suggest that answers like (19b), where the 2W rule is applied, are not necessarily infelicitous in response to questions like (19a), which are a clear indication of a broad focus situation.

- (19) a. Šta se desilo?  
          what REFL happened  
          ‘What happened?’  
      b. Onaj je čovek došao kasno.  
          that AUX.CL man come late  
          ‘That man came late.’

In that sense, we do not agree with the restrictive differentiation given by Diesing et al. (2009) although we think that at least in the Serbian variety, the 2W rule carries some additional pragmatic or semantic cues. Crucially, these additional meanings should not be interpreted in terms of contrastive focus.

The results that we obtained in the study clearly point towards the approach taken by this second group of authors who disagree with the idea that the difference between 2W and 2P is merely sociolinguistic. If that were the case, we would see a clear pattern of difference in the number of instances of the 2W rule among our four groups similar to what we observed with infinitives and DA+present. However, the results show no statistical significance in the way in which clitics are placed within the sample. Virtually all participants opted for the 2P rule. This would not come as a surprise to those who claim that 2W and 2P sentences are different in terms of their syntax (Bošković 2009) and/or semantics and pragmatics (Diesing et al. 2009). The reason why the results are not surprising under the second set of accounts is because our sentences were given without additional contextual information, which would be needed to elicit 2W sentences.



### 4.3 The verb *trebati* ‘need’

Our data show very clearly that the variation between transitive and experiencer *trebati* ‘need’ is confined to the Croatian variety. What is more, ethnicity does not play a major role in the use of this verb. This was shown by the fact that there were no instances of *trebati* as a transitive verb in the ethnically Croatian group from Northern Serbia. In that sense, variation is determined by regional factors.

As transitive *trebati* never occurs in Serbian, the simplest assumption then is that many Croatian speakers have two different lexical items, which some of them may use interchangeably, which is again in line with the approach to variation adopted here. However, before we dismiss the variation with *trebati* as uninteresting and too straightforward, we need to point out the change we note with the modal *trebati* in Serbian. Even though standard/prescriptive grammars and practices go to great lengths to preserve its special status as the only modal which can occur only as an impersonal, the speakers of Serbian show more and more agreeing patterns, whereby *trebati* agrees with the fronted (topicalized or focalized) embedded subject.

- (20) a. **Trebalo je** [da devojke otpevaju tu pesmu].  
 needed.SG.N AUX.3SG DA girls.PL.F sing.3PL that song  
 ‘It was needed / necessary that girls sing this song.’ / ‘Girls should have sung this song.’
- b. Devojke<sub>1</sub> **je trebalo** [da t<sub>1</sub> otpevaju tu pesmu].  
 girls.PL.F AUX.3SG needed.SG.N DA sing.3PL that song  
 ‘Girls should have sung this song.’
- c. Devojke<sub>1</sub> **su trebale** [da t<sub>1</sub> otpevaju tu pesmu].  
 girl.PL.F AUX.PL needed.PL.F DA sing.3PL that song  
 ‘Girls should have sung this song.’

The source of variation in these examples is very interesting because it might be linked to the similarities and differences in the structure of infinitives and DA+present complements discussed in this paper. Namely, the Croatian equivalent of (20a) is (21) where the modal *trebati* has to agree with the subject.

- (21) Djevojke **su trebale** otpjevati tu pjesmu.  
 girl.PL.F AUX.PL needed.PL.F sing.INF that song  
 ‘It was needed / necessary that girls sing this song.’ / ‘Girls should have sung this song.’

The infinitival complement in (21) is incapable of hosting an overt subject and, as Todorović & Wurmbrand (2015) argue, it is quite possible that they do not

even project a syntactic position capable of hosting a subject. Therefore, in Croatian, the subject would have to be base generated with *trebati*, which is why we observe agreement on the modal. On the other hand, DA+present complements always project a SpecvP position, which is sometimes occupied by PRO and sometimes it hosts an overt subject. In (20a), for instance, we find an overt subject with DA+present, hence, the modal *trebati* is impersonal. However, if DA+present and infinitives are in the process of becoming the same structure, as we argued here, the system is forced to accommodate, which is why we are observing the development of a personal use of the previously impersonal modal *trebati*. Based on these facts, we could speculate that the development of a transitive use of the lexical verb *trebati* is linked to this difference in the modal use, but further research is needed to establish this relationship more firmly.

## 5 Conclusions

In conclusion, we have provided empirical evidence that some speakers belonging to the Serbian/Croatian dialect continuum can alternate between infinitives and DA+present constructions without any restrictions even in a written production task. This fact was taken to mean that these are the same underlying structures. Additional syntactic evidence pointing to the same conclusion was also provided. We left open the question of the existence of an active SpecvP position with infinitives as we found some suggestions that the nature of the embedded subject with infinitives and DA+present might be different in certain respects.

A different pattern of variation was found with respect to 2W and 2P clitics. Namely, previous accounts that tie the difference between 2W and 2P clitics to sociolinguistic considerations would predict sharp differences among the four groups of participants in our sample in terms of the use of these two rules for clitic placement. However, such differences were not observed and virtually all participants used the 2P rule exclusively.

Finally, variation in the use of the verb *trebati* as an experiencer verb and as a simple transitive was observed only within the group in Zagreb. No instances of this verb used as a simple transitive have been observed in the groups from Serbia, including the group made up of students with the Croatian ethnic background. In this domain, variation is determined by regional rather than ethnic factors. Also, many speakers who produced sentences with *trebati* as a transitive verb used it as an experiencer verb as well. We have suggested that there are two competing lexical entries for the verb *trebati*, one specified as a transitive verb and the other specified as an experiencer verb, in the mental lexicons of many speakers of Croatian.

The results obtained show a high degree of flexibility in the use of certain syntactic structures like DA+present and infinitives. A significant share of the participants in the study used these structures interchangeably in a controlled production study (i.e. a fixed sociolinguistic context) without any obvious consequences for the semantics and pragmatics of the resulting output. Such a high degree of flexibility is surprising under traditional approaches to syntactic variation where different output structures are expected to arise from different sociolinguistic contexts and/or have different meanings. On the other hand, Adger's (2006) approach creates a much more fluid picture where certain speakers are expected to use different structures interchangeably often without any consequences for the meaning and speaker's decision to use one structure instead of the other is not necessarily triggered by a change in the sociolinguistic context. Since this is precisely what we found with respect to the use of infinitives and DA+present in many speakers of Serbo-Croatian, broader theoretical implications of this study can be found in the fact that it fits into this more fluid picture of syntactic variation proposed by Adger (2006).

## Abbreviations

1	first person	INF	infinitive
3	third person	LR	linear regression
2P	second phrase	NOM	nominative
2W	second word	PASS	passive
ACC	accusative	PL	plural
AUX	auxiliary	PRES	present
CL	clitic	PTCP	participle
DAT	dative	REFL	reflexive
F	feminine	SG	singular
IMP	impersonal		

## Acknowledgements

The work on this paper was supported in part by the project number 178002, entitled *Languages and Cultures in Space and Time*, and project number TR32035, entitled *Development of Dialogue Systems for Serbian and other South Slavic Languages*, both financed by the Government of Serbia.

## References

- Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42 (3). 503–530. DOI:10.1017/S002222670600418x
- Adger, David & Jennifer Smith. 2005. Variation and the Minimalist Program. In Leonie Cornips & Karen Korrigán (eds.), *Syntax and variation: Reconciling the biological and the social*, 149–178. John Benjamins. DOI:10.1075/cilt.265.10adg
- Adger, David & Graeme Trousdale. 2007. Variation in English syntax: theoretical implications. *English Language and Linguistics* 11 (2). 261–278. DOI:10.1017/S1360674307002250
- Anderson, Stephen. 2005. *Aspects of the theory of clitics*. Oxford University Press.
- Bailyn, John. 2010. To what degree are Croatian and Serbian the same language? Evidence from a translation study. *Journal of Slavic Linguistics*. 181–219.
- Belić, Bojan. 2005. The infinitive is difficult to lose: What governs variation of complements in unique control in serbian. *Glossos* 6. 1–20. [https://slaviccenters.duke.edu/sites/slaviccenters.duke.edu/files/media\\_items\\_files/6belic.original.pdf](https://slaviccenters.duke.edu/sites/slaviccenters.duke.edu/files/media_items_files/6belic.original.pdf).
- Bošković, Željko. 2009. On Leo Tolstoy, its structure, Case, left-branch extraction, and prosodic inversion. In Steven Franks, Vrinda Chidambaram & Brian Joseph (eds.), *A linguist's linguist: Studies in south Slavic linguistics in honor of E. Wayles Browne*, 99–122. Slavica Publishers.
- Chomsky, Noam & Howard Lasnik. 1995. The theory of principles and parameters. In Noam Chomsky (ed.), *The Minimalist program*, 506–569. MIT Press. DOI:10.7551/mitpress/9780262527347.003.0001
- Corbett, Greville & Wayles Browne. 2009. Serbo-Croat: Bosnian, Croatian, Montenegrin, Serbian. In Bernard Comrie (ed.), *The world's major languages*, 330–346. London: Routledge.
- Diesing, Molly, Dušica Filipović-Đurđević & Draga Zec. 2009. Clitic placement in Serbian: Corpus and experimental evidence. In Susane Winkler & Sam Featherston (eds.), *The fruits of empirical linguistics*, vol. 2, 59–73. Mouton de Gruyter. DOI:10.1515/9783110216158.59
- Katičić, Radoslav. 1986. *Sintaksa hrvatskoga književnog jezika: Nacrt za gramatiku*. HAZU.
- Kayne, Richard. 2000. *Parameters and Universals*. Oxford University Press.
- Krapova, Ilyana. 1999. Subjunctive complements, null subjects, and case checking in Bulgarian. In István Kenesei (ed.), *Crossing boundaries: Advances in the theory of Central and Eastern European language*, 239–263. DOI:10.1075/cilt.182.14kra

- Kroch, Anthony. 1994. Morphosyntactic variation. In Anthony Kroch & Katharine Beals (eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, vol. 2, 180–201. Chicago Linguistic Society.
- Labov, William. 1972. *Sociolinguistic patterns*. University of Pennsylvania. DOI:10.2307/412715
- Mišeska-Tomić, Olga. 2004. The syntax of the Balkan Slavic future tenses. *Lingua* 114 (4). 517–542. DOI:10.1016/s0024-3841(03)00071-8
- Peti-Stantić, Anita. 2009. Opis, propis i red riječi. In Anita Peti-Stantić (ed.), *Treći hrvatsko-slovenski slavistički skup/tretje hrvaško-slovensko slavistično srečanje*, 85–93. FF Press.
- Ronelle, Alexander. 2006. *Bosnian, croatian, serbian: a grammar*. University of Wisconsin Press.
- Ronelle, Alexander. 2008. Rhythmic structure constituents and clitic placement in bosnian, croatian, serbian. In David Betha (ed.), *American Contributions to the 14th International Congress of Slavists*, vol. 2, 1–20. Slavica Publishers.
- Todorović, Nataša. 2012. *The indicative and subjunctive da-complements in Serbian: A syntactic-semantic approach*. University of Illinois PhD thesis.
- Todorović, Neda & Susanne Wurmbrand. 2015. (In)Finite possibilities of *da*: Restructuring the tense and aspect domains. Presented at the Workshop on Aspect in Embedded Clauses, ZAS Berlin, May 2015.
- Wurmbrand, Susanne. 2003. *Infinitives: Restructuring and clause structure*. Berlin: de Gruyter. DOI:10.1515/9783110908329
- Yu, Kristine. 2008. *The prosody of second position clitics and focus in Zagreb Croatian*. University of California, LA MA thesis. [http://linguistics.ucla.edu/general/matheses/Yu\\_UCLA\\_MA\\_2008.pdf](http://linguistics.ucla.edu/general/matheses/Yu_UCLA_MA_2008.pdf).