STABLE MOOD-TENSE-ASPECT PATTERNS OBSERVED IN THE CLARIN.SI REPOSITORY *

Tanja Samardžić (tanja.samardzic@unige.ch)

1. INTRODUCTION

It is well known, especially to ambitious French learners, that some verbs sometimes take subjunctive mood complements. We can talk about three kinds of verbs illustrated here with 10 examples for each type taken from Grisot et al. (2022):

(1)

Subjunctive complements: *apprécier* 'appreciate', *approuver* 'approve', *attendre* 'wait', *craindre* 'fear', *exiger* 'demand, require', *ordonner* 'order', *permettre* 'allow', *proposer* 'propose, offer', *redouter* 'dread', *souhaiter* 'wish'

Indicative (non-subjunctive) complements: *affirmer* 'state', *constater* 'notice', *croire* 'believe', *déclarer* 'declare', *observer* 'observe', *penser* 'think', *raconter* 'tell', *remarquer* 'remark, note', *trouver* 'find', *voir* 'see'

Both kinds of complements: *accepter* 'accept', *admettre* 'admit', *comprendre* 'understand', *concevoir* 'understand', *imaginer* 'imagine', *nier* 'deny', *rêver* 'dream', *s'assurer* 'make sure', *se plaindre* 'complain', *supposer* 'suppose'

Trying to understand what sets these three groups of verbs apart, we can list a number of lexical and grammatical factors. For example, we can say that the meaning of the complements is more factual in the indicative than in the subjunctive group with the third group being somewhere in between. This opposition has also been formulated in terms of real vs. not real events and emotional vs. cognitive states, but such criteria are hard to relate directly to formal distinctions that can be observed in the behaviour of the complements (Baunaz & Puskás 2014).

Tense and aspect turn out to be observable formal categories interestingly interrelated with the choice of mood. For example, subjunctive mood of the complement is associated with the nonpast tense on the embedded verb in Greek (Giannakidou 2009). Such associations are especially important for investigating the notion of mood in languages that do not mark it overtly. Tense constraints are argued to establish subjunctive values of complements of Slavic

* In the past several years, I have had the pleasure of working with Genoveva on two projects aimed at improving scientific methodology in Linguistics by integrating computational methods into a theoretical framework. This article describes a small study demonstrating how such integration can be implemented to gain a better understanding of the interaction between verb tense, aspect and mood.

GG@G (Generative Grammar in Geneva) 12: 183-194, 2025

verbs too despite the lack of specific morphological marking of the mood on the verb itself: subjunctive complements are associated with the morphological present tense (which might refer to future), while non-subjunctive complements might bear other tenses too (Socanac 2017). When it comes to verb aspect, imperfective is associated with indicative complements, while perfective with subjunctive, which seems to hold across languages (Giannakidou 2009, Socanac 2017).

The interaction of verb aspect with mood is usually discussed at the level of a binary distinction between perfective and imperfective. In Slavic languages, however, verb aspect is famously marked in lexical derivations that can express several categories in a sort of temporal gradation (Samardžić & Merlo 2016), which has to do with other scales too. Forms without prefixes are typically (although not always) imperfective. Once a prefix is attached, the verb becomes perfective and interpreted as more specific (DiSciullo & Slabakova 2005), resultative (Arsenijević 2006) and anchored in time (Ramchand 2008). The prefixed verb can be made imperfective again by adding a suffix in what is known as secondary imperfective. The interpretation of this form remains specific and anchored, but the resultative reading is removed. For example:

(2)

Imperfective	**prič-**a-ti	'tell' (unspecific, not anchored, process)
Perfective	pre-**prič-**a-ti	're-tell' (specific, anchored, result)
Imperfective (secondary)	pre-**prič-**a-va-ti	're-tell' (specific, anchored, process)

The goal of this study is to test empirically theoretical insights about the mood-tense-aspect associations in the complements of the three kinds of verbs listed above and to gain new insights by bringing into the picture a finer granularity of verb aspect. To obtain a more nuanced aspect analysis, I study the Serbian equivalents of the French verbs listed in (1). I extract their instances in a large corpus and collect the information about the properties of their complements, more precisely the form of the embedded verb. I expect the following associations:

(3)

Subjunctive mood	**Indicative** mood
Present tense	**Past** tense (or other)
Perfective aspect	**Imperfective** aspect

At a more fine-grained level of aspect analysis, I look for associations between prefixed verbs, both perfective and imperfective, and mood. The impact of a prefix in an imperfective verb is especially interesting as it might go in two opposite directions: it might underline the process interpretation, making the verb less adapted to a subjunctive context, but its specificity and anchor in time might make it more prone to appear as a subjunctive complement. This study is expected to resolve the direction in which prefixes impact the association between mood and aspect.

2. DATA - THE CLARIN.SI REPOSITORY

The main source of data for this study are observations about language use extracted from large corpora. All of them come from the CLARIN.SI repository,[1] the Slovenian branch of the European CLARIN infrastructure, which specialises in South Slavic languages (although it hosts data in other languages as well). Among other resources, the repository currently hosts over 80 data sets representing Serbian, including various kinds of corpora, lexicons, speech data and processing tools. For the current study, I select two data sets described in more detail in the remainder of this section.

2.1. Serbian web corpus CLASSLA[2]

As one of the many data sets that have been recently created and shared by the CLARIN.SI knowledge centre for South Slavic languages (CLASSLA), the Serbian web corpus (Ljubešić et al. 2024) represents an immense source of examples of language use as its size is more than 2.3 billion words.

What is especially important for linguistic studies is that most of CLASSLA corpora are linguistically annotated at the word level. In our case, this means that each of the over 2.3 billion words is assigned its lemma (citation form) and its morphosyntactic definition (MSD) following the unified standard for all South Slavic languages.[3] This is performed using large language models trained for the task of lemmatisation and MSD tagging (Terčon & Ljubešić 2023a, Terčon & Ljubešić 2023b).

(4)

Word	**Lemma**	**MSD**	Gloss (not available in the corpus)
carica	carica	Ncfsn	empress
naredi	narediti	Vmr3s	orders
da	da	Cs	to
poruše	porušiti	Vmr3p	demolish
taj	taj	Pd-msan	the
hram	hram	Ncmsan	temple

The example (4) illustrates this annotation in a tabular format, where the sentence in the column "Word" should be read vertically. The word-level annotation is added in the other columns. The MSD labels are conventions showing the values of inflectional categories. For instance, 'Ncfsn' stands for a common noun of feminine gender in nominative singular. 'Vmr3s' is a main (non-auxiliary) verb in present tense 3rd person singular, while 'Vmr3p' is the same form in plural. The sequence marked in red is the pattern extracted with methods described below. The corpus can be searched by any of the columns, which is exploited for estimating the mood-tense-aspect associations.

Another important feature of the CLASSLA corpora is their availability for online search via the CLARIN.SI instance of the noSketch corpus query system (Kilgarriff et al. 2014, Rychlý 2007).[4] Although the corpus can be downloaded from the repository and searched locally, manipulating this data size would require much more computational resources than an

[1] https://www.clarin.si/info/about/
[2] http://hdl.handle.net/11356/1931
[3] https://nl.ijs.si/ME/V6/msd/html/msd-hbs.html#msd.V-hbs
[4] https://www.clarin.si/ske/#dashboard?corpname=classlaweb_sr

average personal computer, which makes online search practically the only option for studies of this type.

2.2. HyperVerb database[5]

MSD labels contain the information about verb tenses, but verb aspect is not encoded. To be able to search for verbs of a specific aspect, we need to refer to another resource created by linguists: Database of the Western South Slavic Verb HyperVerb 2.0 -- WeSoSlav (Arsenijević et al. 2024). This database is created by project teams at the Universities of Graz and Nova Gorica and also shared via the CLARIN.SI repository. It contains detailed descriptions of the structure of verbs in Slovenian, Bosnian, Croatian, Montenegrin and Serbian (BCMS) covering 3000 verbs for each language / variety. The size of this data set is relatively small in computational terms, so it is easy to download and manipulate locally.

Among 55 features given for each verb entry, the database contains the binary aspect information (imperfective YES/NO) and also a binary prefix information (has prefix YES/NO). The citation forms extracted from the corpus can thus be matched with the database entries (which are also given in the citation form) to obtain the information about what kind of verb aspect is expressed in the complements of the target matrix verbs.

3. METHODS

For each of the examples taken from Grisot et al. (2022), I first find a Serbian counterpart. It happened that, for each group, one verb had to be left out due to unclear mapping to Serbian, resulting in three groups of 9 target matrix verbs whose instances are searched in the corpus. The information extracted from the corpus is further processed using a custom Python script to obtain the tense and aspect information and calculate percentages. The values obtained from corpus queries are compared to baseline values obtained from both sources of data, the corpus and the database.

3.1. Searching the Serbian web corpus CLASSLA

The noSketch online search interface allows creating advanced queries using the custom corpus query language (CQP).[6] After trying out several options, I created a single query that I used for extracting all verb complementation patterns. For example, the pattern shown in the example (4) is extracted using the following expression:

(5)

```
[lemma="narediti"][]{0,1}[word="da"][]{0,3}[tag="Vm..+"]
```

This expression consists of five elements whose meaning can be described as in (6).

[5] http://hdl.handle.net/11356/1846
[6] https://www.sketchengine.eu/documentation/corpus-querying/

(6)

[lemma="**narediti**"]	:	a verb whose citation form is "narediti"
[]{0,1}	:	0 or 1 other words can appear in between
[word="da"]	:	the complementiser "da"
[]{0,3}	:	0 to 3 other words can appear in between
[tag="Vm..+"]	:	any non-auxiliary verb

In other words, we look for sequences consisting of the target matrix verb, followed by the complementiser, which is in turn followed by any non-auxiliary verb representing the embedded verb. These three main components can be separated by a number of other words.

I made 27 such queries for 27 target matrix verbs with the citation form of the target verb being the only item that changed across these queries. Once all the instances matching a query are found by the system, it is possible to create a frequency list of the found patterns. For each query, I create two frequency lists: one listing the combination of the citation forms (the second column in (4)) and one listing the combinations of the MSD labels (the third column in (4)). I downloaded the first 1000 items in the frequency list corresponding to the top 1000 most frequent patterns. The cut-off of 1000 items is arbitrary corresponding approximately to the point where the frequency of a pattern falls below 0,01% of all hits. The only case where the frequency list contained less than 1000 items was the verb 'opaziti' for French *observer*. In this case, all items were downloaded. The downloaded lists were saved as text files and further processed with a Python script to obtain the counts of mood-aspect and mood-tense pairs, where mood is given by the original French classification and the tense and aspect features are extracted from the corresponding frequency lists.

3.2. Obtaining mood-aspect pairs

To find the aspect information for each verb found as an embedded verb (the expression [tag="Vm..+"]), I wrote a custom Python script that loops over all downloaded citation form patterns and all entries in the HyperVerb database. If an extracted verb is found in the database, the script finds the binary aspect and prefix information and assigns it to the verb. Once the aspect information is known, the script counts the number of cases when the verb was perfective with a prefix, perfective without a prefix, imperfective with a prefix and imperfective without a prefix. In this way, we can tell which target matrix verb is associated with which of these properties of the embedded verb.

3.3. Obtaining tense-aspect pairs

The tense information was extracted from the downloaded MSD patterns with a simple 'Find' function of a text editor that gives the number of hits. I simply looked for the number of 'Vmr' vs. 'Vmp' in the pattern positions that correspond to the embedded verbs (all positions except the first verb, which is the target matrix verb). The string 'Vmr' identifies present tense, while 'Vmp' identifies participles used to form past tenses and conditionals. This means that the latter string identifies all finite non-present forms without differentiating the exact tense. Crossed with the original French classification of the target matrix verbs, the counts of present vs. participle forms can tell us which tense of the embedded verb is associated with which mood type.

3.4. Obtaining baseline values

To assess to what degree the observed associations are characteristic for the target matrix verbs, I compare all the observed scores with baseline values. To obtain the baseline values for verb aspect, I find the number of perfective vs. imperfective verbs in the HyperVerb database. These numbers show whether any of the two aspects is more likely in general, outside of the mood-aspect patterns. In a similar fashion, the baseline preference for prefixation is calculated as the number of prefixed out of all perfective and out of all imperfective verbs respectively.

The tense baseline values are estimated from the corpus by searching for all non-auxiliary verbs (the expression [tag="Vm..+"] outside of the target complementation pattern) and counting how many of them showed up in present tense vs. participle form.

4. RESULTS AND DISCUSSION

Tables 1, 2 and 3 show the values for the three groups of verbs given in (1). Each table contains two main parts: aspect and tense. Each of these two parts contains three main columns showing the proportions of contrasted categories as well as the total counts of distinct unique embedded verbs matched with the corpus query.

To give an example of how to read the tables, we can go through the scores for the verb 'odobriti' for French *approuver.* This verb was found in the corpus complemented by 447 distinct embedded verbs. Out of this number, 248 or 55% of distinct verbs are perfective and 199 or 45% are imperfective. Out of all 248 perfective verbs, 226 or 91% have a prefix and out of 199 imperfective verbs, 66 or 33% have a prefix. Compared to the baseline percentages, we can say that the proportion of perfective verbs in the context of this matrix verb is above the baseline, while the percentage of imperfective verbs is below the baseline. The proportions of prefixed verbs is below the baseline in both perfective and imperfective verbs. Looking into the tense features, we can see that there were 1118 distinct non-auxiliary verbs in the extracted patterns, out of which 1050 were in present tense and only 68 in other tenses. It might seem confusing that the totals are not the same in the cases of aspect and tense. This is due to the fact that tense patterns were not filtered, so they contain items that are not listed in the HyperVerb database. The totals over 1000 are due to the imprecision of the query that sometimes matched two verbs instead of a single embedded verb that was queried. Recall also that this annotation was performed automatically and that there can be mistakes in the assigned labels. The imprecision of the query and of the annotation might bring some noise in the observed patterns, but this does not impact the observed trends.

4.1. Inverse verb aspect preferences, subjunctive vs. indicative and alternating verbs

To see the main trends, we can refer to the mean values in the bottom of each table.

- As expected, the complements that correspond to French **subjunctive** mood tend to have a strong preference for perfective aspect (64% of perfective verbs compared to the baseline of 50%) and a strong preference against imperfective aspect (36% of imperfective verbs compared to the 50% baseline).
- The counterparts of French **indicative** complements show the inverse trend: a preference against perfective (39% vs. 50% baseline) and for imperfective aspect. (61% vs. 50% baseline).
- The counterparts of French **alternating** complements are somewhere in between, but not exactly. They tend to be closer to indicative complements with a preference

against perfective (43% vs. 50% baseline) and for imperfective aspect (57% vs. 50% baseline).

Looking into the within-group variation, we can say that it is somewhat surprisingly small: all verbs in all the three groups follow the observed trends individually as well.

4.2. Present tense preference above the baseline in all three groups

It looks like there is a preference for present tense in general in embedded verbs regardless of the matrix verb. The preference for present is attested in all three groups of complements, but with a different strength. The strongest preference for present tense is in subjunctive counterparts (93% compared to the 67% baseline) and against other tenses (7% against the baseline of 33%). The presence of non-present tenses in this group is so low that it can be attributed to the noise in the data.

The other two groups are again similar with weaker preferences for present tense, but still above the baseline: 71% vs. 67% baseline in the indicative group and 73% vs. 67% baseline in the alternating group. The preference against non-present tenses is below the baseline: 29% vs. 33% baseline in the indicative group and 27% vs. 33% baseline in the alternating group. Here again, the alternating group is in between subjunctive and indicative, but closer to indicative.

The tense pattern raises the question of what makes non-present tenses inconsistent with embedded verbs. We can speculate that embedded contexts require a certain level of temporal underspecificity, which present tense, as a relatively neutral category in Slavic verbs provides rather easily. Other tenses bring more temporal specificity than embedded contexts need, which is why they are avoided in these contexts. The preference for present tense in subjunctive complements is the biggest deviation from the baseline of all the observations. It looks like it is important for the embedded verb to be temporally detached to enable subjunctive reading.

4.3. Prefixation below the baseline in all three groups

The baseline proportion of prefixed verbs is very high (94%) for perfective aspect confirming that prefixation is the main perfectivising mechanism in Serbian verb derivations. The baseline proportion of prefixed forms in imperfective verbs is 45%, confirming their status of a "secondary" imperfective. The proportions observed in all three types of complements (subjunctive, indicative and alternating) are below these baselines. The difference is again in the degree: the contrast between perfective (91% prefixed vs. 94% baseline) and imperfective (27% prefixed vs. 45% baseline) is the strongest in subjunctive complements, while it is less pronounced in the other two groups: 86% prefixed perfective verbs (94% baseline) and 40% prefixed imperfective verbs (45% baseline) in the indicative group; 88% prefixed perfective verbs (94% baseline) and 37% prefixed imperfective verbs (45% baseline) in the alternating group. Once again, the alternating complements are in between the subjunctive and the indicative indicative group, but closer to the indicative.

The prefixation pattern raises two questions. First, why are verb prefixes not welcome in embedded contexts? Second, why would the contrast between perfective and imperfective forms be the strongest in subjunctive complements. The answer to both of these questions might come from the view of prefixes as temporal operators, making the events temporally definite. Since temporal detachment (underspecificity) is consistent with embedded contexts prefixes are avoided, just like non-present tenses. They are especially avoided in subjunctive complements, where temporal detachment is the strongest.

What remains puzzling is the contrast between perfective and imperfective verbs in subjunctive complements. Telling by the smallest difference from the baseline out of all the measures taken in the study (only 3 percentage points), prefixes seem relatively fine in

perfective verbs while the biggest difference compared to a baseline in all mood-aspect pairs is precisely in the category of prefixed imperfective verbs in subjunctive complements (18 percentage points). Prefixed imperfective verbs are the least consistent with a subjunctive context. We can speculate that the availability of a resultative reading in perfective verbs allows the necessary temporal detachment even in the presence of a prefix, but when this reading is not available, the temporal function of prefixes makes them especially inconsistent with subjunctive contexts.

5. CONCLUSION

This study has revealed very stable patterns distinguishing between the same three groups of verbs in French and Serbian. Serbian counterparts of French verbs that take subjunctive complements tend to take perfective verbs in present tense. Serbian counterparts of French verbs that take indicative and alternating complements tend to take imperfective verbs as complements. These findings confirm empirically previous theoretical insights using very large data sets. In addition to the empirical confirmation of the known associations, the study has determined that the alternating complements are more similar to the indicative than to the subjunctive group. This might be a useful observation for further explanations of the alternating complements.

A more fine-grained analysis of the mood-aspect associations showed an interesting parallel between non-present tenses and verb prefixes: both of these categories are avoided in all embedded contexts and especially in subjunctive complements. These findings call for new theoretical accounts potentially relating these categories that belong to different structural levels in Serbian (lexical derivation vs. inflection).

Finally, I would like to underline the methodological side of this study that shows how computationally processed large text data in combination with specific linguistic databases can be used to measure and characterise objectively the notions otherwise intuitively discussed in linguistic theory, in this case the distinction between factual and non-factual or between cognitive vs. emotional. Given the sheer size of the available data and the processing technology, empirical studies such as this one can be expected to become an important source of measurements to be integrated with linguistic theory in the science of natural language.

Subjunctive complements		Aspect					Tense		
		Perfective		Imperfective		All	Present	Particip.	All
		All	Prefix	All	Prefix				
o-**dobr**-i-ti *approuver*	N **%**	248 **55**	226 91	199 **45**	66 33	447 -	1050 **94**	68 **6**	1118 -
o-**ček**-i-va-ti *attendre*	N **%**	325 **65**	292 90	178 **35**	55 31	503 -	1006 **98**	21 **2**	1027 -
strah-o-va-ti *craindre*	N **%**	228 **63**	208 91	134 **37**	53 40	362 -	584 **56**	465 **44**	1049 -
traž-i-ti *exiger*	N **%**	331 **71**	304 92	137 **29**	27 20	468 -	998 **99**	11 **1**	1009 -
na-**red**-i-ti *ordonner*	N **%**	282 **66**	254 90	144 **34**	32 22	426 -	1051 **98**	18 **2**	1069 -
dozvol-i-ti *permettre*	N **%**	335 **59**	307 92	237 **41**	72 30	572 -	1018 **100**	1 **0**	1019 -
pred-**lož**-i-ti *proposer*	N **%**	286 **73**	263 92	105 **27**	21 20	391 -	1013 **99**	11 **1**	1024 -
žel-e-ti *souhaiter*	N **%**	464 **64**	429 92	256 **36**	63 25	720 -	1158 **98**	20 **2**	1178 -
ht-e-ti *vouloir*	N **%**	411 **64**	369 90	228 **36**	44 19	639 -	1185 **95**	57 **5**	1242 -
MEAN	N **%**	323 **64**	295 91	180 **36**	48 27	503 -	1007 **93**	75 **7**	1082 -
BASELINE	N **%**	- **50**	- 94	- **50**	- 45	- -	- **67**	- **33**	- -

Table 1. Aspect and tense features of embedded verbs complementing Serbian counterparts of French verbs taking subjunctive complements.

Indicative complements		Aspect					Tense		
		Perfective		Imperfective		All	Present	Particip.	All
		All	Prefix	All	Prefix				
iz-**jav**-i-ti *affirmer*	N	117	106	193	88	310	699	372	1071
	%	**38**	91	**62**	46	-	**65**	**35**	-
konstat-ov-a-ti *constater*	N	162	145	202	103	364	651	374	1025
	%	**44**	89	**56**	51	-	**63**	**37**	-
ver-ov-a-ti *croire*	N	137	115	171	72	308	758	326	1084
	%	**44**	84	**56**	42	-	**70**	**30**	-
o-**paz**-i-ti *observer*	N	86	71	176	70	262	365	127	492
	%	**33**	83	**67**	40	-	**74**	**26**	-
misl-i-ti *penser*	N	154	134	180	59	334	790	314	1104
	%	**46**	87	**54**	33	-	**72**	**28**	-
prič-a-ti *raconter*	N	147	123	213	65	360	689	400	1089
	%	**41**	84	**59**	30	-	**63**	**37**	-
pri-**met**-i-ti *remarquer*	N	123	101	242	94	365	769	245	1014
	%	**34**	82	**66**	39	-	**76**	**24**	-
u-**vid**-e-ti *trouver*	N	141	124	230	106	371	876	189	1065
	%	**38**	88	**62**	46	-	**82**	**18**	-
vid-e-ti *voir*	N	126	104	242	78	368	778	256	1034
	%	**34**	82	**66**	32	-	**75**	**25**	-
MEAN	N	132	114	205	48	337	708	289	997
	%	**39**	86	**61**	40	-	**71**	**29**	-
BASELINE	N	-	-	-	-	-	-	-	-
	%	**50**	94	**50**	45	-	**67**	**33**	-

Table 2. Aspect and tense features of embedded verbs complementing Serbian counterparts of French verbs taking indicative complements.

Both complements		Aspect					Tense		
		Perfective		Imperfective		All	Present	Particip.	All
		All	Prefix	All	Prefix				
pri-**hvat**-i-ti *affirmer*	N **%**	212 **47**	190 90	224 **53**	82 37	436 -	925 **88**	122 **12**	1047 -
pri-**zna**-ti *admettre*	N **%**	210 **48**	191 91	227 **52**	70 31	437 -	523 **50**	527 **50**	1050 -
s-**hvat**-i-ti *comprendre*	N **%**	131 **47**	109 83	149 **53**	45 30	280 -	896 **81**	206 **19**	1102 -
za-**misl**-i-ti *imaginer*	N **%**	184 **47**	152 83	211 **53**	73 35	395 -	885 **82**	190 **18**	1075 -
neg-ir-a-ti *nier*	N **%**	167 **42**	148 89	229 **58**	96 42	396 -	565 **52**	513 **48**	1078 -
sanj-a-ti *rêver*	N **%**	185 **37**	161 87	314 **63**	92 29	499 -	955 **79**	259 **21**	1214 -
u-**ver**-i-ti se *s'assurer*	N **%**	130 **34**	116 82	203 **66**	96 47	333 -	853 **81**	200 **19**	1053 -
žal-i-ti se *se plaindre*	N **%**	163 **37**	146 90	282 **63**	113 40	445 -	856 **80**	216 **20**	1072 -
pret-po-**stav**-i-ti *supposer*	N **%**	159 **41**	144 91	230 **59**	103 45	389 -	709 **67**	350 **33**	1059 -
MEAN	N **%**	171 **43**	151 88	230 **57**	86 37	401 -	796 **73**	287 **27**	1083 -
BASELINE	N **%**		- 94	- **50**	- 45	- -	- **67**	- **33**	- -

Table 3. Aspect and tense features of embedded verbs complementing Serbian counterparts of French verbs taking alternating complements.

REFERENCES

Arsenijević, B. (2006) *Inner Aspect and Telicity,* PhD thesis, Leiden University.

Arsenijević, B., K. Gomboc Čeh, F. L. Marušič, S. Milosavljević, P. Mišmaš, J. Simić, M. Simonović & R. Žaucer (2024) "Database of the Western South Slavic Verb HyperVerb 2.0 -- WeSoSlav", *Slovenian Language Resource Repository CLARIN.SI*, ISSN 2820-4042, http://hdl.handle.net/11356/1846

Baunaz, L. & G. Puskás (2014) "On Subjunctives and Islandhood", in M.H. Côté & E. Mathieu (eds.) *Variation within and across Romance Languages, Selected Papers from the 41st Linguistic Symposium on Romance Languages (LSRL), Ottawa, 5–7 May 2011*, [CILT 333], John Benjamins, Amsterdam, 233–253. DOI: https://doi.org/10.1075/cilt.333.16bau

DiSciullo, A. M. & R. Slabakova (2005) "Quantification and Aspect", in H. Verkuyl, H. Swart & A. Hout (eds.) *Perspectives on Aspect*, Springer, Dordrecht, 61–80.

Giannakidou, A. (2009) "The Dependency of the Subjunctive Revisited: Temporal Semantics and Polarity", *Lingua* 119(12), 1883–1908. DOI: https://doi.org/10.1016/j.lingua.2008.11.007

Grisot, C., J. Blochowiak & G. Puskás (2022) "Predicting Verbal Mood in Complement Clauses: The Role of the Distributive Patterns, the Tense and the Grammatical Aspect of the Matrix Verb", *Glossa: A Journal of General Linguistics* 7(1), 1–42. DOI: https://doi.org/10.16995/glossa.5798

Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý & V. Suchomel (2014) "The Sketch Engine: Ten Years On", *Lexicography* 2197-4306. 1, 7–36. https://doi.org/10.1007/s40607-014-0009-9

Ljubešić, N., P. Rupnik & T. Kuzman (2024) "Serbian Web Corpus CLASSLA-web.sr 1.0", *Slovenian Language Resource Repository CLARIN.SI*, ISSN 2820-4042, http://hdl.handle.net/11356/1931

Ramchand, G. (2008) "Perfectivity as Aspectual Definiteness", *Lingua* 118, 1690–1715.

Rychlý, P. (2007) *Manatee/Bonito - A Modular Corpus Manager*. RASLAN.

Samardžić, T. & P. Merlo (2016) "Aspect-Based Learning of Event Duration Using Parallel Corpora", in M. Diab, A. Villavicencio, M. Apidianaki, V. Kordoni, A. Korhonen, P. Nakov, M. Stevenson (eds.) *Essays in Lexical Semantics and Computational Lexicography – In Honour of Adam Kilgarriff*, Springer Series Text, Speech, and Language Technology.

Socanac, T. (2017) *Subjunctive Complements in Slavic languages: A Syntax-Semantics Interface Approach,* PhD thesis, University of Geneva.

Terčon, L. & N. Ljubešić (2023a) "The CLASSLA-Stanza Model for Morphosyntactic Annotation of Standard Serbian 2.1", *Slovenian Language Resource Repository CLARIN.SI*, ISSN 2820-4042, http://hdl.handle.net/11356/1831

Terčon, L. & N. Ljubešić (2023b) "The CLASSLA-Stanza Model for Lemmatisation of Standard Serbian 2.1", *Slovenian Language Resource Repository CLARIN.SI*, ISSN 2820-4042, http://hdl.handle.net/11356/1830