

# LINGUISTICA

## LX

**PRISPEVKI O MORFOLOGIJI  
JUŽNOSLOVANSKIH JEZIKOV**

**Topics on the Morphology  
of South Slavic Languages**

Revijo sta ustanovila †Stanko Škerlj in †Milan Grošelj  
Revue fondée par †Stanko Škerlj et †Milan Grošelj

Glavna in odgovorna urednica – Rédactrice en chef  
Martina Ožbot

Številko LX uredili – Responsables du numéro LX  
Christina Manouilidou, Tatjana Marvin, Gašper Ilc, Andrej Stopar

Uredniški svet – Comité de rédaction  
Janez Orešnik, Stojan Bračič, Gregor Perko

Znanstveni svet – Comité scientifique  
Wolfgang U. Dressler (Wien), Martin Maiden (Oxford), Rosanna Sornicola (Napoli),  
Pierre Swiggers (Leuven)

Svetovalni odbor številke LIX – Comité consultatif du numéro LIX  
Božena Cetnarowska, Antonio Fábregas, Valentin Gusev, Sabina Halupka-Rešetar,  
Nikos Koutsoukos, Franc Marušič, Tatjana Marvin, Božena Rozwadowska, Florian Schaefer,  
Marko Simonović, Penka Stateva, Rok Žaucer

Izid revije je finančno podprla  
JAVNA AGENCIJA ZA RAZISKOVALNO DEJAVNOST RS

Publié avec le soutien de  
L'AGENCE NATIONALE SLOVÈNE POUR LA RECHERCHE SCIENTIFIQUE



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. / To delo je ponujeno pod licenco Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna licenca

CIP - Kataložni zapis o publikaciji  
Narodna in univerzitetna knjižnica, Ljubljana

811.163'366(082)

PRISPEVKI o morfologiji južnoslovanskih jezikov = Topics on the morphology of south Slavic languages / [uredili Christina Manouilidou ... et al.]. - Ljubljana : Znanstvena založba Filozofske fakultete = Presses scientifiques de la Faculté des lettres, 2020. - (Linguistica, ISSN 0024-3922 ; 60)

ISBN 978-961-06-0413-6  
1. Vzp. stv. nasl. 2. Manouilidou, Christina  
COBISS.SI-ID 43822851

## VSEBINA

INTRODUCTION .....	5
Boban Arsenijević DEVERBAL NOUNS IN <i>-IE</i> AND THEIR VARIATION ACROSS THE SOUTH SLAVIC AREA .....	7
Marta Petrak DEVELOPMENT OF A PRODUCTIVE DERIVATIONAL PATTERN ON THE BASIS OF LOAN TRANSLATION? THE CASE OF CROATIAN ADJECTIVES FORMED WITH THE PREFIX <i>MEDU-</i> .....	31
Gergana Popova, Andrew Spencer VOLITIONAL MOOD IN SOUTH SLAVIC WITH A FOCUS ON BULGARIAN: A PARADIGMATIC VIEW .....	61
Marko Simonović, Petra Mišmaš √ <i>ov</i> IS IN THE AIR: THE EXTREME MULTIFUNCTIONALITY OF THE SLOVENIAN AFFIX <i>OV</i> .....	83
Marko Simonović CATEGORIES, ROOT COMPLEXES AND DEFAULT STRESS: SLOVENIAN NOMINALIZATIONS REVISITED .....	103
Susanne Wurmbrand, Iva Kovač, Magdalena Lohninger, Caroline Pajančič, Neda Todorović FINITENESS IN SOUTH SLAVIC COMPLEMENT CLAUSES: EVIDENCE FOR AN IMPLICATIONAL FINITENESS UNIVERSAL .....	119



## INTRODUCTION

This special issue of *Linguistica* features a collection of articles dealing with the morphology of South Slavic Languages. The idea behind the volume came from a workshop we organized in the Faculty of Arts, University of Ljubljana, in June 2019, which brought to light various intriguing issues relating to the morphology of the South Slavic languages. This, in combination with the fact that there is no other article collection dedicated to South Slavic morphology, makes the current volume a unique contribution to the topic.

The South Slavic languages belong to the Slavic language family and form a dialectal continuum. They are further divided into West (Bosnian-Croatian-Serbian and Slovenian) and East (Bulgarian and Macedonian) subgroups. Some of them are part of the Balkan *sprachbund* and thus share features with non-Slavic languages spoken in the Balkans. The languages discussed in this volume present a variety of issues relating to morphology and its interaction with other domains of grammar such as phonology and syntax.

The six articles in this special issue examine a range of morphological phenomena either in a cross-linguistic manner across the South Slavic languages (Arsenijević; Wumbrandt et al.) or focus on a specific linguistic phenomenon in a single language, such as Bulgarian (Popova/Spencer), Croatian (Petraček) and Slovenian (Simonovič; Simonovič/Mišmaš). The frameworks used by the authors are diverse and so are their approaches to morphology and its interplay with syntax, semantics and phonology. They either contribute new theoretical insights or provide new descriptive evidence and theoretical argumentation for existing analyses.

The individual contributions are presented in alphabetical order. **Boban Arsenijević's** article deals with *-ie* nominalizations across the South Slavic languages. He proposes a new analysis of the phenomenon, taking into account the prosodic as well as semantic properties of the roots which combine with this suffix. **Marta Petraček** elaborates on the phenomenon of loan affixes in Croatian and provides a diachronic as well as synchronic analysis of the prefix *među-*. Inflectional morphology, and more specifically the expression of volitional mood, is the topic of **Geri Popova's** and **Andrew Spencer's** article; the authors examine the inflectional and analytic forms of expressing volitional mood using the concept of periphrasis, basing their account on data from Bulgarian. **Marko Simonovič** proposes a new analysis of stress assignment in Slovenian deadjectival and deverbal nominalizations. He focuses on the suffixes' influence on stress assignment in Slovenian and proposes, within a Distributive Morphology framework, that the affixes involved in these nominalizations should be treated as roots. Similarly, **Marko Simonovič** and **Petra Mišmaš** combine Distributive Morphology and Optimality Theory to analyze the suffix *-ov-* in Slovenian, providing a unified account of an affix which has traditionally been considered polysemous. Finally, **Susi Wumbrand, Iva Kovač, Magdalena Lohninger, Caroline Pajančič** and **Neda Todorovič** revisit the issue of finiteness in South Slavic complement clauses, a topic which touches upon

morphology, syntax and semantics, and contribute both interesting cross-linguistic data and an innovative analysis.

All papers were subject to rigorous, double-blind review; they received comments from expert peer reviewers and were further read by the editors. We are grateful to the following colleagues who acted as reviewers and contributed to the realization of this volume: Božena Cetnarowska, Antonio Fábregas, Valentin Gusev, Sabina Halupka-Rešetar, Nikos Koutsoukos, Franc Marušič, Tatjana Marvin, Božena Rozwadowska, Florian Schaefer, Marko Simonović, Penka Stateva, and Rok Žaucer.

The Editors  
Ljubljana, November 2020



## DEVERBAL NOUNS IN *-IE* AND THEIR VARIATION ACROSS THE SOUTH SLAVIC AREA

### 1 INTRODUCTION

The suffix *-ie* occurs throughout the Slavic languages as a nominalizing suffix particularly productive with verbal bases. To the best of my knowledge, the analysis of its instantiation in Serbo-Croatian (SC) in Arsenijević (2010) and Simonović and Arsenijević (2014) has been the only attempt so far to capture its general morphophonological and semantic properties. This paper has two goals. The first is to present this analysis and some facts which pose problems for it, as well as to modify it to capture the problematic facts while at the same time providing a more principled account for the prosodic effects of the suffix. The second is to provide an overview of the variation in behavior of the suffix *-ie* across South Slavic languages in light of the category of the base, and to model it based on the modified analysis.

Section 2 introduces the nominalizing suffix *-ie*. Section 3 presents in more detail its properties in SC and section 4 focuses on its prosodic effects. In section 5, the analysis developed in Arsenijević (2010) and Simonović and Arsenijević (2014) is outlined, and some problems that it faces are pointed out. A refinement of the analysis to capture these facts is proposed in section 6. Section 7 gives an overview of the properties of the suffix across South Slavic languages, and section 8 models the variation in terms of the analysis proposed for SC. Section 9 concludes the paper.

### 2 DEVERBAL NOUNS IN *-IE*

In all Slavic languages, the default pattern of derivation of deverbal nouns – the counterpart of the English *-ing* or German *-ung* derivations – involves the suffix *-ie*, added to the uninflected form of the passive participle.

(1)	Czech	Polish	Russian	Old Church Slavonic (OCS)
	lámán- <u>í</u>	łaman- <u>ie</u>	(raz)rušen- <u>ie</u>	pleten- <u>ie</u>
	broken- <u>ie</u>	broken- <u>ie</u>	broken- <u>ie</u>	knotted- <u>ie</u>
	‘breaking’	‘breaking’	‘breaking’	‘knitting’

The suffix takes different phonological and phonetic shapes across the Slavic languages, occurring as: *-je*, *-ie*, *-e*, *-i*. For the sake of uniformity, I refer to it in the text and glosses as the suffix *-ie*.

\* boban.arsenijevic@uni-graz.at

Slavic verbs are marked for grammatical aspect (e.g. Comrie 1976). Each verb bears the aspectual value imperfective or perfective. Several tests can be implemented to verify the aspectual value of the verb. For instance, in the morphological present form, only imperfective verbs may receive the progressive interpretation. Perfective verbs in this form force non-veridical meanings.<sup>1</sup>

- (2) Trenutno čitam / \*pročitam jednu knjigu. *Serbo-Croatian (SC)*<sup>2</sup>  
 currently read.1SG<sup>1p/ Prf</sup> one book  
 ‘I’m currently reading a book.’

Traditional classifications also recognise a third class, termed biaspectual verbs: verbs which are both perfective and imperfective (e.g. Gladney 1982). However, since traditional imperfective verbs normally can be coerced into at least some of the perfective uses, and it is impossible to draw a clear line between biaspectuals, more easily coercable imperfectives and imperfectives which are harder to coerce – the reality of the third class is questionable. Arsenijević (2018) argues that in fact there are only biaspectual and perfective verbs, i.e. that all traditional imperfective verbs are biaspectual, but their perfective interpretations are blocked to a greater or lesser extent by antipresupposition as defined in Percus (2006), due to the availability of their marked perfective counterparts. As they are ambiguous between the two values of aspect, eventual biaspectual verbs do not participate in any contrast relevant for the topic of this paper, and therefore are irrelevant for this discussion. In the interest of simplicity, I am retaining the traditional classification, and considering only the traditional perfective and imperfective verbs.

Across Slavic languages, the deverbal noun (henceforth abbreviated as DN) suffix *-ie* displays different interactions with grammatical aspect, with a trivial selection of perfective bases in Old Church Slavonic (OCS), a preference for perfective basis in Russian, an equal availability of perfective and imperfective bases in Polish and Czech, a slight preference for imperfectives in Slovenian, full productivity with imperfective bases but limited productivity with perfective ones in SC, and exclusive selection of imperfective ones in Bulgarian and Macedonian. The suffix *-ie* also varies in the set of categories with which it combines. While in all Slavic languages it combines with verbs, and is used to derive mass and collective nouns, in some of them it also combines with verb phrases (VP), nouns (N), noun phrases (NP), adjectives (Adj), adjective phrases (AdjP) and preposition phrases (PPs), that is with different subsets of these categories.

- 
- 1 There is, however, variation among Slavic languages in where the boundaries lie between the uses reserved for the perfective verbs, for the imperfective verbs, or available to both. In this section, we abstract away these asymmetries, but later they will be playing an important role in the analysis.
- 2 Henceforth, most examples will be from SC. Therefore, examples from SC will not be specified for language, while all other examples will.



### 3 SUFFIX *-IE* IN SC

In SC, DN<sub>s</sub> in *-ie* from imperfective verbs are productive, compositionally interpreted and denote a homogeneous eventuality – a state or process (Simonović/Arsenijević 2014). When the base verb denotes a state or a process, this is also the eventuality denoted by the DN, as in (3a). A DN from an accomplishment denotes its process subevent or an unbounded sequence of iterations of the accomplishment, as in (3b), and a DN from an achievement denotes its preparatory stage or again an unbounded sequence of iterations of the event, as in (3c).

- |        |   |    |  |    |  |
|--------|---|----|--|----|--|
| (3) a. | spavan-je<br>sleep <sup>lpf</sup> .PASS.PTCP- <i>ie</i><br>'sleeping' | b. | jeden-je<br>eat <sup>lpf</sup> .PASS.PTCP- <i>ie</i><br>'eating' | c. | dosezan-je<br>reach <sup>lpf</sup> .PASS.PTCP- <i>ie</i><br>'reaching' |
|--------|---|----|--|----|--|

Examples in (4) illustrate each of the listed interpretations for the three (types of) verbs.

- (4) a. *a state or process:*

Za vreme spavanja izgubite prosečno 311g težine.  
for time sleeping lose.2HON on\_average 311g weight  
'During sleep, you lose an average of 311g of weight.'

- b. *the process component of an accomplishment:*

Čišćenje dok je dete u kući je kao pranje zuba za vreme  
cleaning while is child in home is like washing teeth for time  
jedenja pakovanja čokolade.  
eating pack.GEN chocolate.GEN  
'Cleaning while a child is in the house is like brushing teeth while eating a bar of chocolate.'

- c. *iterations of an accomplishment:*

Frakcijsko hranjenje – jedenje svaki put u malim obrocima.  
fractional feeding eating every time in small meals  
'Fraction feeding – eating always in small meals.'

- d. *the preparatory stage of an achievement:*

Đivatma zadržava grubo i suptilno telo tokom dosezanja najvišeg  
jivatma preserves rough and subtle body during reaching highest  
stanja kreacije, Kailaša, čistog uma.  
state creation.GEN Kailash.GEN pure.GEN mind.GEN  
'Jivatma preserves a rough and subtle body during the reaching of the highest state of creation, Kailash, the pure mind.'

e. *iterations of an achievement:*

To mu daje osećaj višestrukog dosezanja izgubljenih granica  
 that him gives feeling multiple reaching lost.GEN borders. GEN  
 prostora.  
 space.GEN  
 ‘That gives him the feeling of a multiple reaching of the lost boundaries of  
 space.’

Only an idiosyncratic subset of perfective verbs derive DNs in SC. These DNs take three types of meanings in an unpredictable fashion: the phase transition that introduces quantization into the verb’s semantics (the transition from the culmination not being reached to its being reached), as in (5a), a concrete or abstract object resulting from the eventuality described by the verb, as in (5b) and (5c), respectively.

- |        |   |    |  |    |   |
|--------|---|----|--|----|---|
| (5) a. | zaređen-je<br>consecrate <sup>Prf</sup> .PASS.PTCP- <i>ie</i><br>‘consecration’ | b. | zadebljan-je<br>thicken <sup>Prf</sup> .PASS.PTCP- <i>ie</i><br>‘bump’ | c. | osiguran-je<br>secure <sup>Prf</sup> .PASS.PTCP- <i>ie</i><br>‘insurance’ |
|--------|---|----|--|----|---|

According to Simonović and Arsenijević (2014), the following aspectual asymmetries can be observed on DNs in *-ie* in SC: DNs from imperfectives verbs are productive, compositionally interpreted and prosodically faithful to the base, while among perfective verbs they target an idiosyncratic subset of bases, show a tendency for idiomatic interpretations and bear a long-rising accent on the penultimate syllable, as illustrated in (6).

- |     |               |  |
|-----|---------------|--|
| (6) | imperfective: | koncètrisan+je > koncètrisanje<br>concentrated+ <i>ie</i> ‘concentrating’  |
|     | perfective:   | izjèdnaačen+je > izjednačenje<br>equalized+ <i>ie</i> ‘equalization, phase-transition from<br>a lead to a tie’ (sport) |

The contrasts observed are summarized in (7).

- |     |                  |                         |                                |
|-----|------------------|-------------------------|--------------------------------|
| (7) |                  | Imperfective base       | Perfective base                |
|     | Productivity     | <b>productive</b>       | <b>idiosyncratic</b>           |
|     | Compositionality | <b>compositional</b>    | <b>idiomatic</b>               |
|     | Prosody          | <b>faithful to base</b> | <b>long rising penultimate</b> |

As already pointed out, suffix *-ie* does not only select verbs. With variable productivity, it also combines with nouns and NPs, adjectives and AdjPs, PPs and VPs, as illustrated in (8). In SC, the only category with which the suffix is productive apart from verbs are nouns, with some limited productivity being attested with PPs. A number of *-ie*-nouns from unproductive categories are inherited from older phases in the development of the language – OCS or Old Church Serbian.

(8)	N (productive) kamen-je stone- <i>je</i> 'stones' (collective)	NP (not productive) mal-o-duš-je little-o-soul- <i>je</i> 'despondency'	PP (limited productivity) pri-obal-je by-coast- <i>je</i> 'coastal area'
	VP (not productive) istin-o-ljub-je truth-o-love- <i>je</i> 'love for truth'	Adj (not productive) pošten-je decent- <i>je</i> 'decency'	AdjP (not productive) celo-mudr-ije (archaic) whole-wise- <i>ije</i> 'abstinence/virginity'

The derived category is, however, fixed irrespective of the base: suffix *-ie* universally derives nouns and no other category.<sup>3</sup> Moreover, the nouns it derives – unless lexicalized with an idiosyncratic meaning – tend to denote mereologically homogeneous predicates when it comes to properties of quantity; all the nouns in (8) denote homogeneous predicates. An exception are perfective DNs denoting the respective phase transition to the (holding of the) culmination state, or an object resulting from the event, which can have a quantized interpretation, as illustrated in (9).

- (9) Na dnevnom redu su tri razrešen-j-a.  
on daily order are three discharged-*ie*-PAUC  
'Three discharges are on the meeting agenda.'

#### 4 PROSODIC EFFECTS OF THE SUFFIX *-IE* IN SERBO-CROATIAN

While all Slavic languages do indeed have the nominalizing suffix *-ie*, as stated in the previous section, even just within the Slavic languages, significant variation along two dimensions may be observed. One dimension is its cross-categorial productivity: does it derive nouns from bases of diverse categories, from a subset of them, or only from verbs? The other is its productivity within the verbal category: is it productive only with imperfective verbs, or with perfectives as well?

When it comes to categories, in OCS the suffix is productive across the board: in addition to verbs, it takes nouns and different kinds of NPs, adjectives and AdjPs, PPs, VPs. In Slovenian and SC, besides verbs it only productively selects for nouns, shows restricted productivity with PPs, and, not counting the nominalizations inherited from

3 Considering that the suffix *-ie* derives neuter gender nouns with a stem ending in a palatal segment, and that these nouns have the NSg ending *-e*, strictly speaking only *-ij* is the actual derivational suffix – selecting in the particular case the declension class corresponding to the neuter gender. The same suffix then occurs with the traditional masculine declension class, but only when it selects perfective verbs. It is of limited productivity in this domain and derives nouns denoting episodic quantized occurrences of events (\**(u-)dis*<sup>Prf</sup>-*a-j* 'breath', \**(za-)grli*<sup>Prf</sup>-*a-j* 'hug', where without the perfectivizing suffix, the imperfective verb does not combine with the masculine-selecting *-ij*). This also makes sense considering Arsenijević (2017), who argues that the neuter gender stands for the absence of a syntactic realization of the unit of counting, and masculine for its weak specification. In the interest of simplicity – I ignore this and other potential realizations of the suffix in the present paper.

OCS, does not combine with any other bases. In Macedonian and Bulgarian, again leaving aside loan nominalizations from OCS, it does not take any bases other than verbs. We focus here on the facts relevant for the prosody.

While it is hard to reliably reconstruct the prosodic patterns in OCS, *-ie*-nominalizations inherited from it in South Slavic languages display templatic prosody, typically with properties of default assignment. This is not to say that they all have the same prosodic template, as the default prosodic assignment is sensitive to the category of the base and the phonotactic properties of the components (see Simonović/Arsenijević 2014 for a case study of a default sensitive to the morphological and phonotactic structure).

The inherited DNs, effectively borrowed from OCS, display prosodic prominence on the stem-final syllable – not only in SC, but also in Bulgarian, and trivially so in Macedonian which has a fixed antepenultimate stress (the OCS realization of the suffix is disyllabic). In Bulgarian and Macedonian, the same pattern appears in all other *-ie* nominalizations too, as well as in the phonologically non-adopted ones in SC (each of the Bulgarian examples in (10), with the same prosody, occurs also in SC as a phonologically non-adopted borrowing from OCS).

(10)	<i>SC</i>	<i>Bulgarian</i>		
	V	V	NP	Adj
	otpuštén-je	prosvetlén-je	blagoród-je	pet-o-kniž-je
	indulged- <i>ie</i>	enlightened- <i>ie</i>	noble- <i>ie</i>	five-o-book- <i>ie</i>
	‘indulgence (of sins)’	‘enlightenment’	‘nobility’	‘five-volume set’

In SC, a language with a free lexically specified tone (which then determines the placement of stress), other phonologically adopted *-ie*-nominalizations inherited from OCS show diverse prosodic patterns, all of which display properties of default assignment. When the base is an NP, PP or VP, the suffix *-ie* imposes a short falling accent on the initial syllable, argued to be the main default pattern which is assigned to lexical items lacking lexical prosodic specification (Zec 1999; Simonović/Arsenijević 2014; Simonović 2020). This is illustrated in (11).

(11) a.	NP: pětoknjiižje	b. PP: pričbaalje	c. VP: ĭstinoljublje
	pet-o-knjig-je	pri-obal-je	istin-o-ljub-je
	five-o-book- <i>ie</i>	by-coast- <i>ie</i>	truth-o-love- <i>ie</i>
	‘5-volume set’	‘coastal area’	‘love for truth’

Cases where the base consists of a morpheme lacking any prosody (because it is smaller than a syllable as in (12a), or a clitic as in (12b)) followed by a monosyllabic morpheme are an exception. There, *-ie* nominalizations surface with a rising accent on the initial syllable, as illustrated in (12).

- (12) a. NP: dvòknjiižje                      b. PP: prímoorje  
           dv-o-knjig-je                      pri-mor-je  
           two-o-book-*ie*                    by-sea-*ie*  
           ‘2-volume set’                    ‘coastal area’

Under all available models of SC lexical prosody, these nominalizations are assigned height on the second morpheme of the base – the monosyllabic one. On the assumption that the monosyllabic nature of the second morpheme somehow leave composed clitics outside of the prosodic assignment domain, prosodic assignment in these nominalizations is the same as in those in (11): high tone is assigned to the leftmost syllable in the domain. The only difference is that due to the epenthesis of a vowel or the syllabic status of the clitic morpheme, respectively, the tone may spread one syllable to the left (which is a general rule for SC, see Zec 1999) and thus surface as a raising accent. This is formally represented in (13), where the assignment of the high tone occurs before the emergence of eventual epenthetic vowels, and the latter in some cases does, while in others does not affect the surface prosody.

- (13) a. (pet<sup>H</sup>-knjig-je                      b. (pri<sup>H</sup>-obal-je                      c. (is<sup>H</sup>tin-ljub-je  
           five-book-*ie*                      by-coast-*ie*                      truth-love-*ie*  
           ‘5-volume set’                      ‘coastal area’                      ‘love for truth’
- d. dv-(knjig<sup>H</sup>-je → dv-o<sup>H</sup>-knjig<sup>H</sup>-je                      e. pri-(mor<sup>H</sup>-je → pri<sup>H</sup>-mor<sup>H</sup>-je  
           two-book-*ie*                      by-sea-*ie*  
           ‘2-volume set’                      ‘coastal area’

When the base is a simple count noun, the derivation results in a rising antepenultimate syllable (i.e. the penultimate syllable of the stem), unless the base is monosyllabic, in which case the resulting derivation has a falling accent on the single syllable of the base (the penultimate syllable of the nominalization). This indicates that with a noun as the base, the suffix generally imposes a High tone on the stem-final syllable of the base-noun, which then spreads to the left-adjacent syllable when there is one (Zec 1999).

- (14) a. kàmeen-je → kàmeenje                      b. grân-je → grânje  
           stone-*ie*                      branch-*ie*  
           ‘stones’ (collective)                      ‘branch’ (collective)

Finally, in DNs, i.e. with verbal bases, when the verb is perfective, the prosody of the base plays no role, and the nominalization receives a pattern with a long-rising stem-final syllable, argued to be one of the default prosodic assignment patterns in SC (Simonović/Arsenijević 2014).

- (15) a. pòšten-je > pošćenje  
 decent-*ie*  
 ‘decency’
- b. ispunjen-je > ispunjénje  
 fulfill<sup>Pf</sup>.PASS.PTCP-*ie*  
 ‘fulfillment’

Only when the verb is imperfective does the suffix preserve the prosody of the base, as in (16).

- (16) a. sèčen-je > sèčenje  
 cut<sup>Pf</sup>.PASS.PTCP-*je*  
 ‘cutting’
- b. dozívati > dozívanje  
 call<sup>Pf</sup>.PASS.PTCP-*je*  
 ‘calling’

All in all, *-ie*-nominalizations inherited from OCS which are phonologically adopted, as well as SC DNs from perfective bases, tend to display templatic prosody, in some cases clearly and in others likely emerging from default prosodic assignment. The only class in which lexically specified prosody, in particular the prosody of the base, surfaces on the DN are the imperfective verbs. A tentative generalization emerges that in those cases where the suffix combines with a categorized, compositionally interpreted structure – i.e. with imperfective verbs, the base-prosody is preserved. In all other cases, the surface prosody is default-looking and the suffix combines with a base which is not categorized and/or receives a non-compositional interpretation – because it is borrowed, as in the case of OCS nominalizations, or because it is a root or a root-complex, which is plausibly the case with perfective verbs, nouns (not the entire noun, but typically its root actually derives the *-je* noun, see (14)) and apparent PPs as bases. In support of PP-like bases being root-complexes, the generalization holds that only PPs with morphologically simplex prepositions and complements can be selected by the suffix, and that even though the preposition normally assigns case, the case ending fails to show up in any way in the nominalization, as illustrated in (17). This is to be expected if the base is maximally a pair of roots: that of the preposition and that of the nominal complement.

- (17) a. za            brd-om            za-brd-je  
 behind    hill-INST.N.SG    ‘area behind hill(s)’  
 ‘behind the hill’
- među        rek-ama            među-rek-je  
 between    river-INST.F.PL    ‘area between rivers’  
 ‘between the rivers’
- b. za            pešč-ar-ama        \*za-pešč-ar-je  
 behind      sand-N-INST.N.PL (‘sandpit’)    int. ‘area by sandpits’  
 ‘behind the sandpits’

pri	kotl- <u>in-ama</u>	*pri-kotl-in-je
by	cauldron-N-DAT.F.PL ('ravine')	int. 'area by a ravine'
'by the ravine'		
c. <u>iz-pod</u> brd-a		
from-below	hill-GEN.N.SG	*iz-pod-brd-je
'below the hill'		int. 'area below the hill'
<u>po-kraj</u>	rek-e	*po-kraj-rek-je
over-area	river-GEN.F.SG	int. 'area next to/around a river'
'next to the river'		

## 5 PREVIOUS ANALYSES

### 5.1 Structural Flattening Triggers Prosodic Deletion

The only formal analysis of the prosodic, morphological and semantic properties of the suffix *-ie* in SC proposed so far is that in Arsenijević (2010) and Simonović and Arsenijević (2014). They observe that the opposition between the prosody faithful to the base and the pattern with a long rising penultimate syllable among derived nouns correlates with the opposition between compositional and idiomatic derivation with a range of SC derivational suffixes (besides *-ie*, also *-ost*, *-stvo* and others). Hence, they argue that the long-rising stem-final prosody is the outcome of the default prosodic assignment when the penultimate syllable of the base is long.

They further observe that the suffix *-ie* tends to derive mereologically homogeneous predicates (i.e. mass and collective nouns). Imperfective verbs also denote homogeneous predicates, while predicates denoted by perfective verbs are quantized. The central component of their analysis is that when combined with perfective bases, the homogeneous semantics of the suffix clashes with the quantization of the predicate base, leading to ungrammaticality. No such issue emerges with imperfective verbs. The only way for a passive participle of a perfective verb to combine with the suffix is if its verbal structure which specifies its perfective aspect is erased, effectively forming an adjective, i.e. – an aspectually unspecified base. After all, adjectivization of participles is a productive operation in SC, as well as cross-linguistically. According to this analysis, the structural flattening that lexicalizes the perfective verb into an adjective also erases its lexical prosodic specification, and the derived noun is assigned default prosody.

### 5.2 Marvin (2003): Phasal Spell-Out

An alternative analysis for the prosodic prominence of the suffix in perfective DNs which is applicable to the suffix *-ie* has been proposed by Marvin (2003). In her account, the suffix becomes prosodically prominent when it is a head and is spelled out together with the verbal stem, because they sit in the same phase (Chomsky 2001).

In such a configuration, due to its syntactically stronger status as the phase head, its prosodic specification wins out, and the prosodic specification of the base fails to be realized, as it would yield two prosodically prominent syllables within the phonological word. The suffix cannot bear prominence when it is in a different phase from the base, because the base gets spelled out first, together with its prosodic specification, and by the time the suffix gets spelled out – the prosodic properties of the derived word have already been fixed to realize the specification of the base (the verb in this case).

A possible implementation of this analysis is that when the suffix *-ie* selects an imperfective verb, it selects an AspP which is or contains a phase (even more than one). This means that the verbal base is spelled out separately from the suffix. This explains why the prosody of such DNs is faithful to the base.

When the suffix selects a perfective verb and structural flattening applies, one could in principle argue that it results in the verbal base and the suffix being in the same phase, and hence the suffix surfaces with prosodic prominence.

This latter account faces a severe problem. Without the structural flattening from the other analysis, it requires that with imperfective bases the suffix is not in the same phase as the base, but that it is with perfective roots. Since simplex verbs are typically imperfective (there are not more than a dozen exceptions) and perfectives are typically derived from imperfective verbs, this is an unlikely structure.

Assuming structural flattening, the structure targeted by it does not include the phase head – the adjectival (or participial) suffix *-en/an/t*. Hence structural flattening does not get the suffix *-ie* to the same phase as the verbal base. The analysis hence incorrectly predicts the prosodic prominence of the verbal base. A way out could be to argue that the adjectival/participial suffix, unlike verbal and nominal category heads, does not head a phase, but it is hard to find principled support for such an argument.

### 5.3 Support for Structural Flattening

There is thus an advantage in the analysis relying on structural flattening and prosodic deletion. The view that perfective DNs rely on an operation that turns a structurally complex verb into a root, resulting in the lexical storing of its meaning instead of a compositional interpretation, is supported by the observation that *-ie* nominalizations from perfective verbs are indeed extremely prone to having idiomatic interpretations significantly narrower than those of their base verbs, or metaphorically shifted ones. Their interpretations moreover vary unpredictably between an object and a phase transition to the culmination state, whether abstract or resulting from the described eventuality, as illustrated in (5) above and in (18).

<p>(18) <b>Imperfective, compositional</b>          hlad-jen-je          cold<sup>Impf</sup>-PASS.PTCP-<i>ie</i>          ‘cooling’</p>	<p><b>Perfective, lexicalized (idiosyncratic)</b>          za-hlad-jen-je          for-cold<sup>Prf</sup>-PASS.PTCP-<i>ie</i>          ‘a weather-change to colder’</p>
---	---



<b>Perfective, OCS-loan (idiomatic)</b>	<b>Perfective (compositional): impossible</b>
prikazanje	(*za)-lom-jen-je (za-lom-jen ‘chipped’)
OCS-loan <sup>Prf-<i>ie</i></sup>	for-break <sup>Prf-PASS.PTCP-<i>ie</i></sup>
‘vision’	‘breaking’ (not ‘chipping’)

Additional support for the correspondence between idiomatized meanings and the prosodic pattern comes from the imperfective DNs. Though rare, cases exist where an imperfective verb derives a DN with a shifted interpretation. All such DNs surface with the prosodic template of perfective nominalizations: a long rising accent on the penultimate syllable. The idiomatic, idiosyncratic DN then exists parallel to the productive, compositional one. This is illustrated in (19).

(19)	Base verb		Productive DN		Idiomatic DN
	putòvati <sup>lpf</sup> ‘to travel’	>	putòvaanje ‘travelling’	/	putováanje ‘trip’
	òbrazovati <sup>lpf</sup> ‘to form/educate’	>	òbrazovaanje ‘forming/educating’	/	obrazováanje ‘education’
	òdlikovati <sup>lpf</sup> ‘to decorate’	>	òdlikovaanje ‘decorating’	/	odlikováanje ‘medal’

#### 5.4 Weaknesses of the Analysis

A weak point of the analysis proposed by Arsenijević (2010) and Simonović and Arsenijević (2014) is its (lack of an explicit) account for the deletion of the prosody of the base. It links it to the conversion of the structurally complex verbal participle into a simple adjective by the flattening of its internal structure, but is not specific about the particular operation responsible for the deletion and how it is triggered by the structural flattening.

Moreover, it has been argued (e.g. Trezner 1970) that the set of perfective verbs deriving *-ie* nouns is not entirely idiosyncratic. On a closer look, it turns out that it is almost exclusively the morphologically complex perfective verbs which are not compositionally interpreted from the meaning of an imperfective base verb and a prefix that combine with the suffix (which we shall term *primary perfective verbs*). There are two big classes of such verbs: idiomatic perfectives, as in (20a), where the verb involves an intransparent root which is idiosyncratic for this verb (observe the absence of a corresponding passive participle even without the prefix), and those borrowed from OCS, as in (20b).<sup>4</sup>

4 OCS verbs have a hybrid status regarding complexity. They involve prefixes which are always homophonous with SC counterparts (even when OCS had a different realization for a prefix, it is adapted to its SC counterpart) and simple verbs which may or may not have SC counterparts, and for this reason they are identified as morphologically complex. However, their meaning is typically not compositional in this regard, so they are recognized as very similar to idiomatic verbs.

- (20) a. pri-sp-et-je                      \*sp-et                      b. prestavi-en-je  
       by-??<sup>Prf</sup>-PASS.PTCP-*ie*    ??-PASS.PTCP                      pass\_over<sup>Prf</sup><sub>OCS</sub>-PASS.PTCP-*ie*  
       ‘arrival’    ‘death’ (the event)

This puts the structural flattening in a different perspective. If the perfective verbs deriving DNs are already non-compositional, then the structural flattening is not a last resort operation licensing the derivation of a DN, but rather occurs as an independent property of the verb that makes its participle available to the DN derivational pattern.

There seem nevertheless to be additional factors involved, as not all primary perfectives (those derived from morphologically simple imperfectives) derive DNs. Some of them do not sound perfect, such as the example in (21a), but become much more acceptable in an appropriate context, as in (22a). Others, like the one in (21b), remain bad across salient contexts – see (22b).

- (21) a. ruš-en                                      ruš-en-je  
       destroy<sup>lprf</sup>-PASS.PTCP                      destroy<sup>lprf</sup>-PASS.PTCP-*ie*  
       ‘destroyed’                                      ‘destroying’
- pre-ruš-en                                      ?pre-ruš-en-je  
       over-dress<sup>Prf</sup>-PASS.PTCP                      over-dress<sup>Prf</sup>-PASS.PTCP-*ie*  
       ‘disguised’                                      ‘getting disguised’
- b. građ-en                                      građ-en-je  
       build<sup>lprf</sup>-PASS.PTCP-*ie*                      build<sup>lprf</sup>-PASS.PTCP-*ie*  
       ‘built’    ‘building’
- na-građ-en                                      ??na-građ-en-je  
       on-build<sup>Prf</sup>-PASS.PTCP                      on-build<sup>Prf</sup>-PASS.PTCP-*ie*  
       ‘awarded’                                      ‘getting awarded’

- (22) a. Njegovim    prerušenjem    u    boga    kiše  
       his.INST    disguising.INST    in    god    rain.GEN  
       počinje    prolećni    ritual    plodnosti.  
       begins    spring    ritual    fertility.GEN  
       ‘With him disguising into the god of rain begins the spring ritual of fertility.’
- b. ?Njegovim    nagrađenjem    za    životno    delo                      počinje  
       his.INST    awarding.INST    for    life    achievement.GEN    begins  
       festival    vina.  
       festival    wine.GEN  
       ‘With him receiving the life achievement award begins the festival of wine.’

Note that both sentences in (22) would be fine with the respective imperfective DNs, but with a different interpretation. In that case, it would be the process or the preparatory stage, and not the phase transition into the culmination state that marks the beginning of the ritual, i.e. of the festival, respectively.

The problem that emerges with this change of perspective is that it implies that the idiomatic semantics does not emerge through the structural flattening of the participle required for nominalization as previously argued, but rather presents an independent property of the verb. The participle then cannot be aspectually different from the verb, i.e. it is perfective and quantized.

Moreover, primary perfectives quite regularly derive secondary imperfectives with which they establish aspectual pairs: the two verbs semantically minimally differ in the aspectual value, as in (23). The view in Arsenijević (2010) that *-ie*, which derives mereologically homogeneous predicates, takes imperfective bases because they also denote homogeneous predicates, predicts that here, too, *-ie* will combine with the secondary imperfective only, which will block the DN derived from the perfective verb. However, such perfective verbs still derive DNs on their idiomatic interpretation, while the DNs from their imperfective counterparts preserve the compositional interpretation.

(23)	po-stavi-ti	:	po-stavi-a-ti		u-blaži-ti	:	u-blaži-va-ti
	over-put <sup>Prf</sup> -INF		over-put-IPF <sup>IPf</sup> -INF		in-milden <sup>Prf</sup> -INF		in-milden-IPF <sup>IPf</sup> -INF
	‘place’ (Prf)		‘place’ (IpF)		‘milden’ (Prf)		‘milden’ (IpF)
	postavljenje		postavljanje		ublaženje		ublaživanje
	‘appointment’		‘placing’		‘mildening’		‘mildening’
			‘appointment’		(of measures)		(any)

If idiomatization is not a tool to enable perfective DNs but an independent process, then why are perfective *-ie* DNs available in the first place? Why is the *-ie* DN of verbs like in (23) not derived only from the secondary imperfective? In particular since there are also DN suffixes restricted to perfective verbs, such as *-aj* and *-a* illustrated in (24): why does a perfective verb combine with the homogeneous, rather than a quantized DN suffix?

(24)	u-dis-aj		po_kuš-aj		iz_daj-a		pro_val-a
	in-breathe <sup>Prf</sup> -aj		try <sup>Prf</sup> -aj		betray <sup>Prf</sup> -a		break_in <sup>Prf</sup> -a
	‘inbreath’		‘try <sub>N</sub> ’		‘betrayal’		‘burglary’

## 6 A MODIFIED ANALYSIS

In this section, I present a modification of the analysis presented in section 5, necessary to capture the empirical insights from subsection 5.4, as well as an extension based on Simonović (2019), which gives a principled and independently motivated account for the prosodic regularities.

## 6.1 Suffix *-ie* in SC May Select Roots – It Is Only Incompatible with a Quantized AspP

A small modification of the analysis from Arsenijević (2010) and Simonović and Arsenijević (2014) suffices to account for the observed facts. Idiomatic and borrowed primary perfective verbs are regular perfectives, yet their perfectivity is not structurally derived but stored as part of the lexical semantics of their respective roots. They are thus effectively non-derived perfectives. Their intransparent semantics triggers their storing in the lexicon as roots. For instance, the intransparent meaning of the morphological complex *u-stolič* ‘in-chair’ (meaning *enthroned* rather than *into (a) chair*) triggers the storing of the complex in the lexicon as a root instead of a compositional interpretation in a verbal structural sequence, as illustrated in (25).

$$(25) \quad [\text{AspP } u\text{-} [\text{vP } [x] \text{-i } [[y] \sqrt{\text{USTOLIČ}} [[y] u [z]]]]] \rightarrow \sqrt{\text{USTOLIČ}}$$

$\lambda x \lambda y \lambda e [\exists e1, e2. \text{cause}(e, e1, e2) \ \& \ \text{act}(e1, x) \ \& \ \text{in}(e2, y, z) \ \& \ \text{chair}(z)]$

When these verbs take verbal inflection, they project the perfective structure with the respective aspectual head filled by a null element (with possible head-movement into it by the root and the material it picks up). However, when they build DN<sub>s</sub>, the roots are directly selected by the adjectival suffix *-en/an/t*. This structure is then selected by the DN suffix *-ie*. This is illustrated for the two verbs from (20), one idiomatized and one borrowed primary perfective, in (26) below.

$$(26) \quad \begin{array}{ll} \text{a. } [\text{TP } \text{-m } [\text{AspP } \text{-}\emptyset \text{ } [\text{vP } \text{-i } [\sqrt{\text{USTOLIČ}}]]]] & \text{b. } [\text{TP } \text{-m } [\text{AspP } \text{-}\emptyset \text{ } [\text{vP } \text{-i } [\sqrt{\text{OSVEŠT}}]]]] \\ \text{ustolič-i-m} & \text{osvešt-a-m} \\ \text{enthroned-}\theta\text{-} \text{vowel-1sg} & \text{sanctified-}\theta\text{-} \text{vowel-1sg} \\ \text{'I enthroned'} & \text{'I sanctified'} \end{array}$$
  

$$\begin{array}{ll} [\text{nP } \text{-ie } [\text{adjP } \text{-n } [\sqrt{\text{USTOLIČ}}]]] & [\text{nP } \text{-ie } [\text{adjP } \text{-n } [\sqrt{\text{OSVEŠT}}]]] \\ \sqrt{\text{USTOLIČ}}\text{-en-je} & \sqrt{\text{OSVEŠT}}\text{-an-je} \\ \text{in-throne}^{\text{Prf}}\text{-ADJ-}ie & \text{sanctify}^{\text{Prf}}_{\text{OCS}}\text{-ADJ-}ie \\ \text{'enthronement'} & \text{'sanctification'} \end{array}$$

Rather than triggering the flattening of the perfective verbal structure to delete its quantized semantics, the suffix *-ie* simply selects for the suffix *-en/an/t* which in turn does not select a quantized AspP but a root (for a general overview of the suffix *-en/an/t*, see Simonović/Arsenijević 2014). Suffix *-en/an/t* selecting a root is a proper target on this view.

This means that, in principle, for every verb the suffix may select either the fully-fledged passive participle – as long as its AspP is not specified as quantized, or just the root combined with the adjectivizing suffix.

With imperfective verbs, both options are available, the participle being preferred due to its compositional interpretation. Therefore, imperfective verbs productively derive compositional DNs, and idiosyncratically root-based ones, i.e. only where an idiomatic interpretation is pragmatically supported, as in (19).

With perfective verbs, the option with the participial base is not available, as it involves an AspP specified for quantization. Idiomatic primary perfective verbs act as roots available as legitimate bases for DNs.

But on this view, structural flattening cannot be responsible for the prosodic effects. If it had such effects, they should be visible already on the primary perfective verbs.

## 6.2 Radical Cores

The prosodic contrast between the prosody faithful to the compositional base and the default-looking prosody of DNs with root-bases receives a neat explanation from the model proposed by Lowenstamm (2014), in particular in the version developed for the South Slavic data in Simonović (2019). Lowenstamm argues within the framework of Distributed Morphology (DM, Halle/Marantz 1993) that (at least some) derivational morphemes are roots. He views these roots as semantically light, and dissociates them from the functional effects that they coincide with – such as the category change. Category change does not come from these suffixes, but from an independent phonologically null category head.

Simonović, also in the framework of DM, radicalizes Lowenstamm’s view that some derivational affixes are roots: he takes all phonological material to come from roots and all functional material to be phonologically empty. Traditional derivational suffixes are chunks of structure consisting of an empty category head and a light root (the one surfacing as the suffix), which are stored in the Encyclopedia. Simonović provides additional empirical support for this view, and argues that in Slovenian, configurations with at least one functional head between every two roots results in resolved prosody (i.e. surface prosody matching the lexical specification of one of the components), illustrated in (27a), and configurations with radical cores (structural segments consisting of two or more structurally adjacent roots without any intervening functional items) trigger default prosody, as in (27b).

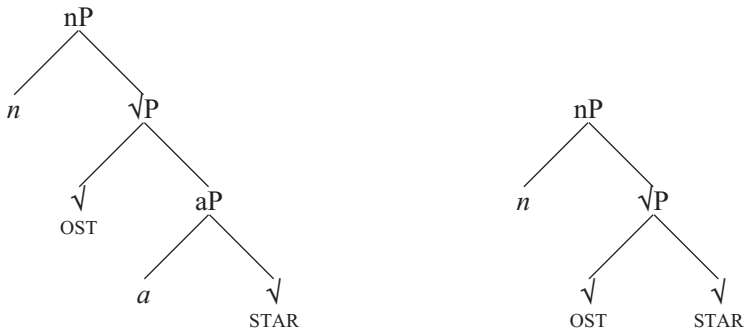


(27) a. Resolved prosody (typically, lex. prosody of  $v_1$  surfaces)

b. Default prosody

Simonović argues, based on minimal pairs such as those with the deadjectival nominalizing suffix *-ost* in (28), that prosody faithful to the base surfaces in compositionally interpreted derivations because they involve no radical cores, while their idiomatically interpreted counterparts carry the default stress, which is stem-final in Slovenian.

(28)                      stár ‘old’ + *ost*                      Slovenian  
                                  stárost ‘oldness’                      staróst ‘old age’

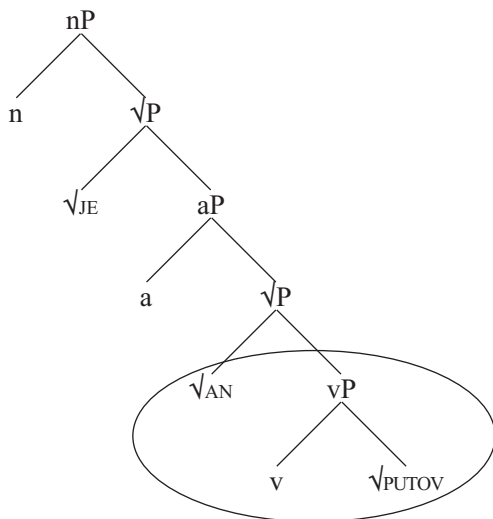


### 6.3 Applying the Modified Analysis to SC DNs

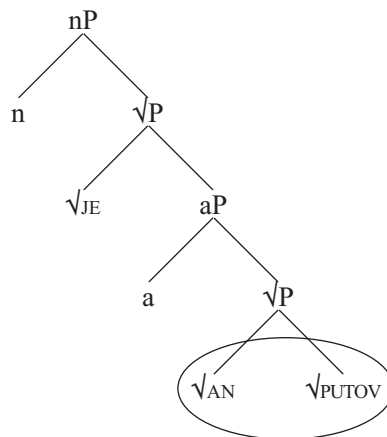
With this view adopted, the default stress does not require deletion – it is straightforwardly predicted from the root status of the base. Consider (29), where the structural contrast from Simonović is applied to the minimal pair of two DNs derived from the same verb. One DN is derived from the participle, i.e. from the combination of the adjectivizing chunk (i.e. [<sub>aP</sub> a [√<sub>AN</sub> ... ]]) and the verb (the vP), as represented in (29a). The other is formed from the combination of the adjectivizing chunk with a root complex, as represented in (29b).<sup>5</sup>

5 At the merger of two roots, one of them semantically projects, by projecting its ontological class. The emerging asymmetry can be stipulated to suffice to drive the LF interpretation, as well as the phonological linearization, but fails to properly feed the prosodic processing, thus resulting in the default prosody assignment.

(29) a. A compositional DN,  
no radical cores



b. An idiomatic DN:  
a radical core



This analysis both fits the data better, considering that the perfective verbs with corresponding DNs are independently idiomatized and that imperfective verbs also may derive idiomatic DNs with the default prosody, and also provides a principled account for the assignment of default prosody.

## 7 SUFFIX *-IE* ACROSS THE SOUTH SLAVIC LANGUAGES

South Slavic languages show a great deal of variation in the productivity and selection properties of the suffix *-ie*.

### 7.1 OCS

In OCS, the suffix is highly promiscuous and productive. It productively combines with PPs, nouns and NPs (involving numerals or attributive adjectives), verbs and VPs, adjectives and AdjPs, as illustrated in (30) and (31).

(30)	N	NP	PP	<i>OCS</i>
	kamen- <i>ie</i>	pet-o-knjig- <i>ie</i>	za-bræg- <i>ie</i>	
	stone- <i>ie</i>	five-o-book- <i>ie</i>	by-hill- <i>ie</i>	
	'stones' (collective)	'5-volume set'	'outback'	
	VP	Adj	AdjP	
	hrist-o-ljub- <i>ie</i>	blažen- <i>ie</i>	čelo-mqdr- <i>ie</i>	
	Christ-o-love- <i>ie</i>	blessed- <i>ie</i>	whole-wise- <i>ie</i>	
	'love for Christ'	'bliss'	'abstinence/virginity'	

Both imperfective and perfective verbs in OCS derive DNs, but perfective DNs are about ten times more frequent. A vast majority of imperfective nominalizations, if not all, are of the idiomatized type, in the sense that they do not denote the process component of the described event, but typically its product. This is illustrated by the DNs deriving from imperfective verbs in (31), where rather than denoting, respectively, the states of knowing (*znanie*), of not bearing (*neterpenie*), the preparatory states or processes of blinking (*mgnovenie*) and reading (*čtenie*), each of them denotes an individual associated with the event-kind described by the base.

(31)

im-an-ie	zn-an-ie	ne-terp-en-ie	mgnov-en-ie	čt-en-ie	<i>OCS</i>
have-PASS.PTCP- <i>ie</i>	know-PASS.PTCP- <i>ie</i>	not-bear-PASS.PTCP- <i>ie</i>	blink-PASS.PTCP- <i>ie</i>	read-PASS.PTCP- <i>ie</i>	
‘property’	‘knowledge’	‘impatience’	‘moment’	‘text’	

Finally, the passive participle in *-en/an/t* in OCS is universally perfective, irrespective of the grammatical aspect of the verb, which means that ultimately all the bases of *-ie* DNs in OCS are perfective. In other words, the OCS suffix *-ie* effectively selects no imperfective (i.e. homogeneous) bases whatsoever. This is probably also the explanation for the preference that the suffix shows for perfective verbs: these verbs are more likely to be used in the perfective passive participle form.

## 7.2 Macedonian and Bulgarian

In Macedonian and Bulgarian, the picture is the exact opposite. The suffix only combines with imperfective verbs, and in this domain it is fully productive, see the examples in (32a). Native perfective verbs do not derive DNs – see (32b). Both languages have a limited number of perfective DNs inherited from OCS or Russian Church Slavonic (RCS), but these are clearly marked as borrowed, rather than derived words. In both languages, their ending differs segmentally from the native DN suffix, and in Bulgarian they are also characterized by the antepenultimate stress inherited from the source language. The DNs borrowed from OCS and RCS, include a small number of imperfective idiomatic ones. Three such Bulgarian DNs of OCS origin, two perfective and one imperfective, are given in (32c).

(32)

a. bǎǎa-n-e	kára-n-e	mísle-n-e	<i>Bulgarian</i>
run <sup>Impf</sup> -PASS.PTCP- <i>ie</i>	drive <sup>Impf</sup> -PASS.PTCP- <i>ie</i>	think <sup>Impf</sup> -PASS.PTCP- <i>ie</i>	
‘running’	‘driving’	‘thinking’	
b. *iz-bǎǎa-n-e	*pre-kára-n-e	*iz-mísle-n-e	
away-run <sup>Prf</sup> -PASS.PTCP- <i>ie</i>	over-drive <sup>Impf</sup> -PASS.PTCP- <i>ie</i>	out-think <sup>Impf</sup> -PASS.PTCP- <i>ie</i>	
‘running’	‘driving’	‘thinking’	



c. <i>nastroé-n-ie</i> direct <sup>Prf</sup> -PASS.PTCP- <i>ie</i> 'mood'	<i>opravdá-e-ie</i> justify <sup>Prf</sup> -PASS.PTCP- <i>ie</i> 'justification'	<i>ži-t-ie</i> live <sup>Prf</sup> -PASS.PTCP- <i>ie</i> 'life' (the literary genre)
---	--	--

### 7.3 Serbo-Croatian

As already discussed, suffix *-ie* is fully productive with imperfective verbs, and limited to (a subset of) those perfectives which are not compositionally derived. It also productively combines with nouns, deriving collectives, and with limited productivity with PPs, deriving mass nouns.

### 7.4 Slovenian

When it comes to non-verbal bases, suffix *-ie* in Slovenian behaves like SC: it productively combines with nouns, and with limited productivity also with PPs. When it comes to verbs, a difference can be observed. Like in Bulgarian, Macedonian and SC, the suffix is productive with imperfective verbs (although for some bases blocked by other DN suffixes, in particular *-va*). Unlike in these other three contemporary South Slavic languages, in Slovenian, the suffix *-ie* is also relatively productive with perfective verbs, and certainly not limited to the loan vocabulary and/or the idiomatic perfective verbs. This is illustrated in (33).

(33)		<i>Slovenian</i>
<b>PP (limited productivity)</b>	<b>N (productive)</b>	<b>Perfective V (productive)</b>
pod-kolen-je under-knee- <i>je</i> 'knee pit (area)'	list-je leaf- <i>je</i> 'leaves' (collective)	raz-cvet-an-je away-flower <sup>Prf</sup> -PASS.PTCP- <i>je</i> 'blooming'

Some of the perfective DNs give a first impression of degradation, but they most often improve in an adequate context, and can easily be found in colloquial use. In any case, there is a large number of compositionally interpreted ones like *razcvetanje* 'blooming' above, *podkurjenje* 'lighting up', *začaranje* 'enchantment', which would be out in Macedonian, Bulgarian and SC.

Finally, Slovenian DNs differ from all others in systematically preserving the prosody of the base, irrespective of its grammatical aspect. Consider the examples in (34). Both stand for a pair of verbs, one with the prefix in brackets and the other without it (*prečiščenje* and *čiščenje*, *začudenje* and *čudenje*). Prefixed versions are perfective and prefixless imperfective. In both variants, each of the two DNs has the same prosody – that inherited from the base.

(34)	(pre)-číst-iti over <sup>Prf</sup> -clean <sup>lpf</sup> -INF 'clean (up)'	(za-)čúd-iti for <sup>Prf</sup> -wonder <sup>lpf</sup> - <i>iti</i> '(begin to) wonder'	<i>Slovenian</i>
------	--	---	------------------

(pre)-číst-jen-je ((pre)čišćenje)	(za-)čúđ-en-je
over <sup>Prf</sup> -clean <sup>IpF</sup> -PASS.PTCP- <i>je</i>	for <sup>Prf</sup> -wonder <sup>IpF</sup> -PASS.PTCP- <i>je</i>
‘cleaning (up)’	‘wondering / beginning to wonder’

Simonović (2019) points out a small number of exceptions with a stem-final stress.

(35)		<i>Slovenian</i>
mišljen > mišljén-je, *mišljen-je	vprašan > vprašán-je, *vprašan-je	
‘thought’ ‘opinion, thinking’	‘asked’ ‘question’	

All these exceptional DNs involve idiomatic perfective bases.

## 8 MODELLING VARIATION

The main observations based on the overview of the suffix *-ie* in South Slavic are the following:

1. There is a substantial difference between the behavior of the suffix *-ie* in OCS and in the contemporary South Slavic languages: in OCS it is quite unselective, and combines with several types of expressions that are unavailable to it in the four living varieties: with different kinds of NPs, with adjectives and AdjPs, and with VPs. In Slovenian and SC, the suffix is productive with nouns and shows restricted productivity in the PP domain, while in Bulgarian and Macedonian it is limited to imperfective verbs.
2. It takes only perfective bases in OCS, both perfective and imperfective in SC and Slovenian – with a strong preference for imperfective bases in SC, and a mild one in Slovenian, and only imperfective bases in Bulgarian and Macedonian.

I propose the following account. The suffix *-ie* is the default nominalizer in OCS, while in the living South Slavic varieties it has grammaticalized to carry homogeneous semantics. In OCS, it takes (homogeneous and) quantized bases (perfective verbs, numeral-noun expressions), in contrast to the living varieties, where it avoids the latter. Among the contemporary South Slavic languages, only in Bulgarian and Macedonian, it is further restricted to the verbal category. It prefers perfective bases in OCS because they are more likely to derive the form that it selects: the perfective passive participle. In SC and Slovenian, it combines with verbal bases, and – under pragmatic licensing – also with roots.

Where both are available and semantically equivalent, the suffix prefers compositional bases to roots. The reason why it shows a decent degree of productivity with perfective bases in Slovenian is that Slovenian verbal aspect has significantly bleached the quantized semantics of the perfective verbs. This observation has been made independently (Dickey 2003), and can be supported by many contrasts, such as for instance the availability of perfective verbs in different homogeneous contexts, in which they cannot be used in other Slavic languages. Consider the contrasts in the use of perfective verbs for the present performative semantics (36).

- (36) *Slovenian* SC Dickey (2003:196–197)  
 Obljubim, da tega ne bom več naredil. : \*Obečam da to više neću raditi.  
 promise<sup>Prf.1SG...</sup> promise<sup>Prf.1SG...</sup>  
 ‘I promise I won’t do that ever again.’  
 Priznamo. Strah nas je. : \*Priznamo. Strah nas je.  
 admit<sup>Prf.1PL...</sup> admit<sup>Prf.1PL...</sup>  
 ‘We admit it. We’re afraid.’

We arrive at a picture where suffix *-ie* has been grammaticalized from a general nominalizer in OCS into one restricted in terms of the category and properties of quantity of the base in all the living South Slavic languages. In Bulgarian and Macedonian it is specified to select only verbal bases, in Slovenian and SC – roots are valid targets as well, and in OCS its domain is categorially unrestricted. The weakened quantized semantics of perfectives in Slovenian licenses a more liberal combination of the suffix with verbal bases regarding the grammatical aspect.

## 9 CONCLUSION

This paper has focused on two different perspectives of *-ie*-DNs: an analysis of DN in contemporary SC and the variation across South Slavic languages. After presenting the main properties of *-ie* nominalizations, a modified analysis is proposed to fit the observation that the perfective verbs which derive DN are all already idiomatically interpreted. It is further combined with the Lowenstammian model of prosodic assignment from Simonović (2019) in order to capture the prosodic effects of the suffix in a principled way.

An overview of the general situation in South Slavic is provided, which is then modelled in terms of the analysis proposed for SC with a refinement varying the specification of the suffix across the target languages. The empirical contrasts observed are shown to follow from two aspects of the specification of the suffix: whether it is sensitive to the structurally realized quantization, and whether it is restricted to verbs.

Its good fit to the South Slavic data lends additional support to Lowenstamm’s (2014) and Simonović’s (2019) view of derivational suffixes as roots, and to the analysis by the latter where radical cores, i.e. structurally adjacent roots, result in default prosody.

## References

- ARSENIJEVIĆ, Boban (2010) “On two types of deadjectival nominalisation in Serbian.” *Suvremena lingvistika* 36/70, 129–145. <https://hrcaak.srce.hr/62322>
- ARSENIJEVIĆ, Boban (2017) “Gender, like classifiers, specifies the type of partition: Evidence from Serbo-Croatian.” In: J. Kantarovich/T. Truong/O. Xherija (eds), *Proceedings of the 52<sup>nd</sup> Annual Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 21–37.

- ARSENIJEVIĆ, Boban (2018) “Serbo-Croatian suffix -va and the illusion of the imperfective specification.” Paper presented at the conference *Ambigo: Workshop on Ambiguity – Theory, Development, and Processing, University of Göttingen on 4–6 July 2018*. Göttingen: University of Göttingen.
- COMRIE, Bernard (1976) *Aspect*. Cambridge: Cambridge University Press.
- DICKEY, Stephen M. (2003) “Verbal Aspect in Slovene.” In: J. Orešnik/D. F. Reindl (eds), *Sprachtypologie und Universalienforschung* 56/3. Berlin: Akademie Verlag, 182–207.
- GLADNEY, Frank Y. (1982) “Biaspectual Verbs and the Syntax of Aspect in Russian.” *The Slavic and East European Journal* 26/2, 202–215. <https://doi.org/10.2307/308089>
- HALLE, Morris/Alec MARANTZ (1993) “Distributed morphology and the pieces of inflection.” In K. Hale/S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111–176.
- LOWENSTAMM, Jean (2014) “Derivational affixes as roots: Phasal spell-out meets English stress shift.” In: A. Alexiadou/H. Borer/F. Schafer (eds), *The syntax of roots and the roots of syntax*. Oxford University Press: Oxford, 230–259.
- MARVIN, Tatjana (2003) Topics in the stress and syntax of words. Doctoral dissertation. Cambridge, MA: Massachusetts Institute of Technology.
- PERCUS, Orin (2006) “Antipresuppositions.” In: A. Ueyama (ed), *Theoretical and Empirical Studies of Reference and Anaphora: Toward the establishment of generative grammar as an empirical science, Report of the Grant-in-Aid for Scientific Research (B), Project No. 15320052*. Fukuoka: Japan Society for the Promotion of Science, 52–73.
- SIMONOVIĆ, Marko/Boban ARSENIJEVIĆ (2014) “Regular and honorary membership: On two kinds of deverbal nouns in Serbo-Croatian.” *Lingue e linguaggio* 13/2, 185–210. <https://doi.org/10.1418/78407>
- SIMONOVIĆ, Marko (2019) “Radical core, prosody preservation and default stress: Revisiting Slovenian nominalisations.” Manuscript. Graz: University of Graz.
- SIMONOVIĆ, Marko (2020) “Tonal spans in Neo-Štokavian.” Manuscript. Graz/Nova Gorica: University of Graz/University of Nova Gorica.
- TREZNER, Tomislav (1970) “Glagolske imenice na -nje.” *Jezik: časopis za kulturu hrvatskoga književnog jezika* 18/2, 50–54.
- ZEC, Draga (1999) “Footed tones and tonal feet: rhythmic constituency in a pitch-accent language.” *Phonology* 16, 225–264. <https://doi.org/10.1017/S0952675799003759>

Abstract  
DEVERBAL NOUNS IN *-IE* AND THEIR VARIATION  
ACROSS THE SOUTH SLAVIC AREA

The paper proposes an analysis of the correlation between the semantic and prosodic properties of the suffix *-ie* and of its variation across the South Slavic languages. Empirical facts about the suffix are outlined, and previous analyses are presented and confronted with empirical and theoretical problems. A slight modification of the analysis of Arsenijević (2010) and Simonović and Arsenijević (2014) enables its combination with the model from Simonović (2019, 2020). The combined analysis neatly accounts for the facts. A model of the variation in the behaviour of the suffix across South Slavic languages is formulated in terms of the modified analysis.

**Keywords:** deverbals nominalizations, radical cores, South Slavic, grammatical aspect, default prosody

Povzetek  
IZGLAGOLSKI SAMOSTALNIKI Z MORFEMOM *-IE* IN  
NJIHOVE RAZLIČICE V JUŽNOSLOVANSKIH JEZIKIH

Prispevek predstavi analizo korelacije med pomenskimi in prozodičnimi lastnostmi morfema *-ie* in njegovih različic v južnoslovanskih jezikih. V članku orišemo empirična dejstva, predstavimo dosedanje razčlembе ter navedemo z njimi povezane empirične in teoretične zagate. Razprava nato preoblikuje analizo Arsenijevića (2010) in Simonovića in Arsenijevića (2014), tako da postane združljiva z modelom Simonovića (2019, 2020). Delno preoblikovana razčlembа lepo pojasni obravnavane podatke; prispevek z njeno pomočjo uvede model, ki razloži razlike v vedenju obravnavanega morfema v južnoslovanskih jezikih.

**Ključne besede:** izglagolske nominalizacije, korenski skupek, južnoslovanski jeziki, slovnčni vid, privzeta prozodija





## DEVELOPMENT OF A PRODUCTIVE DERIVATIONAL PATTERN ON THE BASIS OF LOAN TRANSLATION? THE CASE OF CROATIAN ADJECTIVES FORMED WITH THE PREFIX *MEĐU*-<sup>1</sup>

### 1 INTRODUCTION

This paper deals with Croatian adjectives containing the prefix *među-* ‘inter-’, the majority of which are derived on the basis of the [*među*-N-Suff]<sub>Adj</sub> pattern. Such adjectives are a result of the simultaneous addition of a prefix and a suffix to a nominal base, as in the examples *među-grad-ski* ‘intercity’<sup>2</sup> (<*grad* ‘city’) or *među-zub-ni* (<*zub* ‘tooth’) ‘interdental’ (Babić 2002: 445). While this word-formation pattern is productive in contemporary Croatian, diachronic data seem to demonstrate that it accounted for just a few adjectives only about a hundred years ago.

More precisely, when one analyses older available dictionaries, such as the Dictionary of the Croatian or Serbian Language (also known as the Academy’s Dictionary, Budmani/Maretić 1904–1910), one finds only several such adjectives, most of which are either hapax legomena or recent inventions formed as equivalents of Latin or German terms. In line with that important fact, one could arguably presume that the [*među*-N-Suff]<sub>Adj</sub> pattern did not exist in the Croatian language, but that it emerged only sometime in the 19<sup>th</sup> century, first as a result of loan translations of Latin and German terms. The only adjective listed in the Academy’s Dictionary that is not claimed to be a loan translation (calque) is *međusoban* ‘mutual’ (<*među* ‘between’ + *se*(*be*) ‘oneself’). However, this adjective, according to the information provided, only appears in the works of several authors in mid- or late 19<sup>th</sup> century. These facts seem to underpin the conclusion that it was probably in the 19<sup>th</sup> century that *među-* prefixed adjectives first appeared in Croatian.

The next available lexicographic source, Benešić’s dictionary (1957), comprising Croatian literary works published from the beginning of the 19<sup>th</sup> century until the

\* mpetrak@ffzg.hr

- 1 An earlier version of this paper was presented in the form of a poster at the Mediterranean Morphology Meeting (MMM12) in Ljubljana in June 2019. I would like to thank two anonymous reviewers for their numerous comments and suggestions, which were almost entirely included in this revised version of the paper.
- 2 All the English equivalents in the paper, unless otherwise indicated, were provided from the Merriam Webster online dictionary available at: <https://www.merriam-webster.com/> (accessed in August 2020). The writing rules, especially those related to hyphenation, were preserved as they appear in the dictionary.

1940s, lists only three *među-* adjectives: *međunarodan* ‘international’, *međusatni* ‘occurring between two (school) classes’ and *međuzvezdani* ‘interstellar’. This fact, while based on literary works only, still points to the conclusion that *među-* adjectives were not that common in mid-20<sup>th</sup> century Croatian. Newer dictionaries, however, list more *među-* adjectives; thus in *Veliki rječnik hrvatskoga standardnog jezika* (2015) (the *Large Dictionary of Croatian Standard Language*, henceforth: VRH 2015), the largest dictionary of the Croatian standard language, one finds 37 such adjectives.

In order to complement the analysis of the lexicographic works, a part of the study presented in this paper also included an analysis of three digital corpora: Riznica, a corpus containing literary and other written sources from the second half of the 19<sup>th</sup> century to this day, HNK (Hrvatski nacionalni korpus), the Croatian national corpus comprising texts written from 1990 onwards, and hrWaC, a web corpus and the largest extant corpus of the Croatian language. The corpora results seem to corroborate the lexicographic ones. More precisely, in the Riznica corpus, as few as four *među-* adjectives appear between late 19<sup>th</sup> century and the 1970s, and they start to become more numerous from the 2000s onwards. In the HNK, there are 51 *među-* adjectives with 10 or more occurrences, and in hrWaC there are as many as 134 *među-* adjectives with 10+ occurrences. These facts point to the conclusion that, over time, the [*među-N-Suff*]<sub>Adj</sub> pattern has started to produce more and more adjectives, some of which have not yet been listed in dictionaries. The absence of certain *među-* adjectives found in the hrWaC corpus from dictionaries might, firstly, be a result of the fact that some of them are specialized terms which one would not expect to find in general language dictionaries such as the VRH, and secondly, of the fact that it takes some time for dictionaries, especially paperback ones, such as the VRH, to list new lexemes that have already existed in the language for some time.

The central question that we explore in this paper is whether the [*među-N-Suff*]<sub>Adj</sub> pattern appeared as a result of loan translation.

Even though lexicographic works such as the Academy’s dictionary seem to suggest that *među-* adjectives appeared as loan translations of Latin and German terms in the 19<sup>th</sup> century, Croatian borrowing literature says little about phenomena related to the borrowing of derivational morphemes, with rare exceptions (cf. Kovačec 1967: 101–103; Dabo-Denegri 2007: 44). When it comes to the borrowing of morphemes generally, it is usually separated into borrowing of inflectional morphemes, which was considered by some prominent linguists as virtually impossible (e.g. Meillet 1926; Jakobson 1938), and borrowing of derivational morphemes, which some authors consider to be more susceptible to borrowing (e.g. Weinreich 1953). Today, there is general agreement among linguists that inflectional borrowing is a relatively rare phenomenon, while derivational morphemes are more easily borrowed (Gardani, Arkadiev and Amiridze 2015: 9).

In this paper, we hypothesize that adjectives formed according to the [*među-N-Suff*]<sub>Adj</sub> pattern first appeared in Croatian as a result of loan translation under Latin and German influence, and that the pattern was subsequently reinforced under strong English influence. Even though this hypothesis cannot be entirely confirmed, once the



[*među*-N-Suff]<sub>Adj</sub> pattern appeared in Croatian, it probably continued to exist to some extent, as supported by evidence of the Riznica corpus and Benešić's dictionary. However, it was only in the mid- and late 20<sup>th</sup> century that it really became productive, as attested to by the VRH dictionary, as well as the HNK and hrWaC corpora. That period coincides with a large English influence on Croatian (Filipović 1990; Samardžija 2002; Turk 2013), so it could be assumed that the influence of English contributed to the reinforcement of the [*među*-N-Suff]<sub>Adj</sub> pattern, and to the creation of a larger number of new *među*- adjectives.

The rest of the article is organized as follows: Section 2 provides an overview of language borrowing phenomena, with a special focus on morphology, especially that which is derivational, and the difference between MAT and PAT borrowing (Sakel 2007). Section 3 positions the Croatian language in the context of language borrowing, emphasizing especially the role of loan translations (calques). In Section 4, we provide more details about the derivation of adjectives in Croatian. Section 5 is dedicated to the methodology we have used to study the question of whether *među*- adjectives appeared in Croatian as a result of loan translation. Our methodology is a combination of analyses conducted on available lexicographic works and digital corpora. Section 6 presents the results of our analysis. It shows that, as predicted at the beginning, *među*- adjectives first emerged as a result of loan translation, and have become more numerous with time, with the highest number of occurrences recorded in contemporary Croatian, both according to lexicographic and corpus data. The majority of adjectives are formed via the [*među*-N-Suff]<sub>Adj</sub> pattern, and the *među*- prefix in these adjectives behaves like a polysemous morpheme whose meanings form a radial structure, with the prototypical (spatial) meaning at the centre (cf. Lakoff 1987). In Section 7 we provide some concluding remarks, reaffirming our hypothesis that *među*- adjectives have entered the Croatian language as a result of loan translations of Latin and German terms, and could subsequently have become more numerous under English influence. This is also corroborated by data from the Etymological Dictionary of the Croatian Language (2016).

## 2 BRIEF OVERVIEW OF LANGUAGE BORROWING PHENOMENA

Linguistic change stemming from language contact could be said to be an almost universal phenomenon occurring in both existing and extinct languages (cf. Grant 2019: 1). Another well-known fact is that language contact can affect the participating languages on a number of levels, and most prominently in the lexicon (e.g. Thomason 2001: 10; Turk 2013; Poplack 2018: 1; Grant 2019: 1). In this paper, however, attention is turned to a different aspect of language contact, morphological borrowing. The issue has been much less explored than lexical borrowing, but today is “far from being unknown” (Grant 2019: 17). Moreover, recent years have seen an increasing scholarly interest in the field of morphological borrowing (Gardani, Arkadiev and Amiridze 2015: 1).

There are various phenomena that enter the field of morphological borrowing. They are generally divided into two major types: borrowing of inflectional morphemes (e.g. Weinreich 1963: 31–33; Kovačec 1967; Minayeva 2003), and borrowing of derivational morphemes (e.g. Della Volpe 1997; Coghill 2015).

Sakel (2007) distinguishes two basic types of borrowing between languages: matter borrowing (MAT) and pattern borrowing (PAT). MAT-borrowing refers to taking over both morphological material and its phonological form, while PAT-borrowing refers to replicating the organization and grammatical or semantic meaning without borrowing the form itself. This second type results from loanshifts or calques (Sakel 2007: 16). In other words, PAT-borrowing refers to the situation in which a recipient language uses its own morphological elements and organizes them in a way that resembles the structure of the source language (cf. Gardani, Arkadiev and Amiridze 2015: 3).

This paper focuses on a particular type of PAT derivation: loan translation of complex words consisting of an affix (a prefix), a (nominal) base and a suffix. A general claim in the literature on affix borrowing is that affixes are most commonly borrowed indirectly, as part of complex loanwords (Weinreich 1963: 31–32). In such a case, a number of complex loanwords enter a recipient language, and only subsequently can its speakers analyse these words into their constituent parts and eventually start using a foreign affix with native bases (cf. Seifart 2015: 512). In direct borrowing, on the other hand, an affix is recognized by the speakers of a recipient language and immediately used with native bases (*ibid.*). For Seifart (2015), the crucial difference between the two types of affix borrowing lies in the question of whether speakers of a recipient language understand the source language. If they do, it can lead to a situation of direct borrowing.

Loan translations or calques are a specific type of borrowing. In the case of complex words, a number of authors distinguish between an “entire calque”, or the reproduction of a complex foreign word element by element (*loan translation proper*, Weinreich 1963: 51; Martinet 1980: 170), and a “partial calque”, or the loan translation of only parts of words (*loan rendition*, Weinreich 1963: 51; *calque approximatif*, Martinet 1980: 170). When applied to words formed with affixes, examples of the first case would be loan translations of both an affix and a base (e.g. French *pré-élargissement* ‘pre-enlargement’ -> Polish *przedrozszerzenie*, Trajder 2007: 140), and of the second one loan translation of the affix only (e.g. French *supercommissaire* ‘super-commissioner’ -> Polish *superkomisarz*) (*ibid.*).

In the examples analysed in this paper, i.e. in Croatian *među-* prefixed adjectives, entire calque is at work due to the fact that both the prefix (*među-*), the base and (usually) the suffix are borrowed and translated, i.e. expressed with Croatian linguistic material, such as in the following example: Latin *internationalis* (*inter-* ‘inter- + *natio* ‘nation’ + *-alis* ‘adjectival suffix’ = ‘international’) > Croatian *međunarodan* ‘international’ (*među-* ‘inter-’ + *narod* ‘people’ + *-an* ‘suffix’). Before proceeding with the analysis of adjectives, some details need to be provided about major linguistic influences on Croatian, as well as about the formation of Croatian adjectives in general.

### 3 CROATIAN IN THE CONTEXT OF LANGUAGE CONTACT

At its very beginnings, Croatian already came into contact with several substrate languages, and had lasting contact with Latin as the language of Western Christianity, administration and education (Samardžija 2002: 61). By the end of the Middle Ages, it had

also established contacts with a number of neighbouring languages: Italian, Hungarian, German and later on Turkish, all of which exerted influence primarily on the lexical level, at times very strongly<sup>3</sup> (Samardžija 2002: 61–62). Latin loanwords from the areas of Christianity and philosophy, but also related to administration, law and new inventions, were for a long time the most numerous ones in Croatian (Samardžija 2002: 63).

Calques became a regular phenomenon in the Croatian lexicon starting from the second half of the 16<sup>th</sup> century, and were especially related to the publication of first larger dictionaries (Samardžija 2002: 63). Older Croatian lexicographic works were usually bi- or even multi-lingual, and their source language was usually a foreign one, mostly Latin (Gostl 1995). Faced with numerous gaps on the Croatian side, lexicographers were oftentimes forced to invent equivalents themselves, which resulted in a large number of neologisms, calques, etc. (Samardžija 2002: 64).

At the time of the industrial revolution, a considerable number of new technical and scientific terms were formed on the basis of classical languages (Latin and Greek), which are sometimes called Europeisms (Croatian *europaizmi*) due to their presence in a number of modern European languages (Samardžija 2002: 65). At the end of the 19<sup>th</sup> century, Croatian borrowed a number of internationalisms through its contacts with German and Italian (Samardžija 2002: 65). It needs to be emphasized that Latin was the official language in continental Croatia until 1847, followed by German until 1860 (Samardžija 2002: 66). When Croatian finally became the official language, it lacked functional diversity. It therefore saw numerous additions in the second half of the 19<sup>th</sup> century, during which time two prominent lexicographers played a key role: Šulek in continental Croatia and Parčić in littoral Croatia (Samardžija 2002: 66–67). They both agreed on providing Croatized words for all notions where it was possible (Samardžija 2002: 67), which left the language once again with a substantial portion of calques.

In recent history, English is undoubtedly the language that has exercised by far the strongest influence on Croatian (Samardžija 2002: 72). Numerous Anglo-American elements have spread into Croatian owing primarily to the media, and have entered the language at a quick pace (Samardžija 2002: 72). Moreover, Turk (2013: 159) claims that in the second half of the 20<sup>th</sup> century Croatian was “inundated” by loanwords from English. English influence on the Croatian language has occurred both overtly, in the acceptance and adaptation of English lexemes into the lexicon, and covertly, as loan translations, which are “really numerous” (Muhvić-Dimanovski 1992: 94), and can be found on virtually all language levels (Margić Drljača 2009).

#### 4 ADJECTIVAL WORD-FORMATION PROCESSES IN CROATIAN

Having given an overview of language contact phenomena relevant for this paper, this section shall provide some more details about the formation of adjectives in Croatian. The major word-formation processes on the basis of which Croatian adjectives are formed are the following: suffixation (e.g. *glazba* ‘music’ + *-en* ‘suffix’ > *glazben* ‘musical’), prefixation (e.g. *ne-* ‘un-’ + *služben* ‘official’ > *neslužben* ‘unofficial’),

3 Multinational states such as the Austro-Hungarian Monarchy, which Croatia was a part of, were an important factor that contributed to language contacts (Turk 2013: 15).

prefix-suffix combination (e.g. *izvan-* ‘out-’ + *brak* ‘marriage’ + *-ni* ‘suffix’ > *izvanbračni* ‘extramarital’), compound-suffix combination (e.g. *hladan* ‘cold’ + *krv* ‘blood’ + *-(a)n* ‘suffix’ > *hladnokrvan* ‘cold-blooded’) and compounding (e.g. *vatra* ‘fire’ + *otporan* ‘resistant’ > *vatrootporan* ‘fireproof’) (Babić 2002: 381; 445; 459; 463; 472–473). The word-formation process that accounts for the majority of adjectives is suffixation (Babić 2002: 381).

When it comes specifically to the question of word-formation with the prefix *među-*, the author of the most comprehensive manual on Croatian word-formation, Babić (2002: 445; 461; 468; 473), claims that it participates in the following four types of adjective formation:

- 1) prefix-suffix combination of relational adjectives (e.g. *međugradski* ‘intercity’, *međuzubni* ‘interdental’, *međunarodni* ‘international’, etc.),
- 2) prefix-suffix formation of descriptive adjectives (1 example: *međusobni* ‘mutual’),
- 3) prefix-suffix formation of descriptive adjectives with zero suffix (1 example: *međuvremen* ‘intertime’),
- 4) formation of descriptive adjectives through prefixation (1 example: *međuzavisan* ‘interdependent’).

From the abovementioned list, it can firstly be concluded that all types of adjective formation but the first one – prefix-suffix combination of relational adjectives – are rather unproductive and of very limited scope, because they are all used to form a single adjective, according to Babić (2002). Unlike these three processes, prefix-suffix combination results in a number of relational adjectives. Secondly, it can also be concluded from the aforementioned facts that the *među-* prefix is productive in the formation of relational, and not descriptive adjectives.

Prefix-suffix combination, or the formation of new lexemes through the simultaneous addition of a prefix and a suffix, is also called parasynthetic formation or parasynthesis (e.g. Serrano-Dolader 2015; Iacobini 2020). The term *parasynthesis* is mostly used today to refer to Romance verbs formed from adjectival or nominal bases (e.g. French *embarquer* ‘to load, board’ < *em-* ‘in’ + *barque* ‘boat’ *-er* ‘infinitive ending’) (Serrano-Dolader 2015: 524), but some authors also use it to refer to nouns and adjectives formed through the addition of a prefix and a suffix to a base (Serrano-Dolader 2015; Iacobini 2020). It is important to emphasize that, in order for a formation to be considered a case of parasynthesis, many authors argue that there should not be an attested “intermediate stage”: thus, in the above French example, there are no words such as *\*barquer* or *\*embarque*. Authors writing from a generative point of view explain that requirement on the basis of the binary branching hypothesis, which specifies that only one word-formation process can apply at a time (cf. Serrano-Dolader 2015). In other words, they reject the possibility of ternary structures for parasynthetic derivations (*ibid.*).<sup>4</sup>

---

4 For more details on the treatment of parasynthesis in linguistic literature, see Serrano-Dolader (2015).

When applied to Croatian adjectives formed according to the [*među*-N-Suff]<sub>Adj</sub> pattern, however, this principle does not seem to work formally because a simple adjective can be found in the language for every parasynthetic one, such as in the following examples: *gradski* ‘urban’  $\diamond$  *međugradski* ‘intercity’; *državni* ‘state’  $\diamond$  *međudržavni* ‘interstate’; *zvjezdani* ‘stellar’  $\diamond$  *međuzvjezdani* ‘interstellar’, etc. Still, all these adjectives are claimed to be parasynthetic or prefix-suffix formations by Babić (2002) due to the fact that their meaning cannot be construed as the sum of the prefix and an adjective: for instance, *međugradski* ‘intercity’ does not mean ‘occurring between what pertains to the city’ (\**među*- ‘inter-’ + *gradski* ‘pertaining to the city’), but its meaning is ‘relative to what is between cities’, thus *među*- ‘inter-’ + *grad* ‘city’ + *-ski* ‘suffix’. In this paper, such adjectives are considered parasynthetic formations.

The fact that Babić (2002) enumerates a number of adjectives formed through prefix-suffix combination with *među*- points to the conclusion that adjectives formed according to this process are fairly present and regular in contemporary Croatian. What Babić (2002) omits to specify, however, is, firstly, how productive the derivational pattern [*među*-N-Suff]<sub>Adj</sub> is, and secondly and more interestingly, how it emerged in Croatian.<sup>5</sup> It is therefore the goal of this paper to shed some light on the history of the formation of *među*- prefixed adjectives and to explore their productivity in present-day Croatian language.

## 5 METHODOLOGY

In order to explore the emergence of *među*- prefixed adjectives in Croatian and the productivity of the patterns according to which they are created, available lexicographic works and corpora were consulted. More precisely, three dictionaries were used: the Academy’s Dictionary (Budmani/Maretić 1904–1910), Benešić’s dictionary (1957), and VRH (2015). Brief explanations shall be given as to why the three dictionaries mentioned were chosen for the analysis presented in this paper.

The Academy’s Dictionary is a rich historical dictionary of Croats, Serbs, Bosnian-Herzegovinian Muslims and Montenegrins that provides information from the earliest linguistic sources in the 12<sup>th</sup> century up until works of the 19<sup>th</sup> century. In the period when it was written, it was considered that these ethnic groups spoke a single language called Croato-Serbian or Serbo-Croatian. Importantly for this paper, it is also a terminological dictionary, as well as a dictionary of foreign words and loanwords.

The exact title of Benešić’s dictionary (1957) is *Rječnik hrvatskoga književnoga jezika od preporoda do I. G. Kovačića* (*Dictionary of the Croatian Literary Language from the National Revival until I. G. Kovačić*). Its intention was to be a dictionary of contemporary Croatian literary language as a collection of quotes from the most excellent Croatian writers who published between the very beginning of the 19<sup>th</sup> century until the 1940s (Nikolić-Hoyt 2010: 63–64). It was chosen due to the fact that it covers the “middle” period between early 20<sup>th</sup> century and the 1940s. One of Benešić’s goals in compiling his dictionary was to revise and modernize the data found in earlier

5 The second question was outside Babić’s (2002) scope because he wrote a synchronic word-formation manual.

dictionaries, for instance by leaving out words that were no longer used in the Croatian literary language, and by introducing those that were used by Croatian modern authors (Nikolić-Hoyt 2010: 62). It should therefore serve as a good illustration of the Croatian lexicon from early 19<sup>th</sup> century to mid-20<sup>th</sup> century.

Finally, the VRH dictionary is the largest and most recent dictionary of the Croatian standard language (Slišković 2016: 244). It is based on older relevant lexicographic works, manuals, specialized dictionaries and digital corpora.

After an analysis of the mentioned lexicographic works, three major digital corpora were also consulted: Riznica, HNK and hrWaC. The Riznica corpus, compiled by the Institute of Croatian Language and Linguistics, contains literary and other written sources from the second half of the 19<sup>th</sup> century to this day.<sup>6</sup> The corpus contains 100 million tokens (Brozović Rončević and Čavar 2012). Due to its specificities, Riznica was searched via the *među*.\* standard regular expression, providing all words starting with the graphic sequence *među*.

The HNK, 3.0 beta version, contains more than 2.3 billion words (Tadić 2009). Much larger than Riznica, it is a balanced and representative corpus<sup>7</sup> of standard contemporary Croatian, which contains a certain amount of fiction (such as magazines, newspapers, books, diaries, novels, etc.), fiction, etc., in line with text typology standards (Tadić 2002: 442). The HNK corpus was also searched via the *među*.\* standard regular expression, providing all words starting with the graphic sequence *među*. The results obtained were then organized through the *Frequency – Lemma* function, providing a list of *među*- words with their number of occurrences in the corpus. The results were manually checked.

The hrWaC (Ljubešić and Erjavec 2011) is a web corpus whose 2.2 version was crawled in 2014 from the .hr domain, so it provides us with data about very recent Croatian language usage. It is the largest extant Croatian corpus, with 1.4 billion tokens. Adjectives entering the analysis were extracted from the corpus via the following CQL order:

```
[word="među.*"] containing [tag="A.*"]
```

The order searches for all words beginning with the sequence *među* and bearing the PoS mark “A”, i.e. adjective. After that, using the option *Frequency – Lemma*, all the obtained results were organized according to their frequency of appearance in the

---

6 The Riznica corpus comprises the following: fundamental works of Croatian literature, popular works, scientific works and university manuals from different domains, elementary and high school manuals, translations by prominent Croatian translators, daily, weekly and monthly newspapers available online, books from the pre-standard period of the Croatian language, see <http://riznica.ihjj.hr/dokumentacija/index.hr.html>.

7 See [http://filip.ffzg.hr/cgi-bin/run.cgi/corp\\_info?corpname=HNK\\_v30](http://filip.ffzg.hr/cgi-bin/run.cgi/corp_info?corpname=HNK_v30). The corpus consists of two components: 1) written contemporary Croatian texts, dating from 1990 onwards, and 2) the so-called text archive, comprising various genres published before or after 1990, such as classical Croatian authors, but also chatroom discussions, etc. (Tadić 2002: 443).

corpus. For the purposes of this paper, only adjectives with  $\text{freq} \geq 10$ , which enables the finding of both high- and lower frequency lexemes, were taken into consideration. The corpus data was checked manually in order to eliminate noise such as adjectives formed via other word-formation processes (e.g. *međunarodno-pravni* ‘related to international law’), typos (*međusubni*, *međunardni*), etc., leaving a final list of 134 adjectives. This figure itself already suggest that adjectives formed with the prefix *među-* are fairly numerous in contemporary Croatian, and that the derivational pattern is a rather productive one.

## 6 RESULTS AND ANALYSIS

This section first presents the results of the lexicographic analysis, followed by corpus results.

### 6.1 Older Dictionaries

#### 6.1.1 Academy’s Dictionary

Table 1 presents the *među-* adjectives in the Academy’s Dictionary (1904–1910).

Table 1: Adjectives formed with *među-* in the Academy’s Dictionary (1904–1910).

	Adjective	Etymology (as specified by the Academy’s Dictionary)	Comment (from the Academy’s Dictionary)
1	<i>međudnevički</i> ‘related to <i>međudneвица</i> ’	Derived through suffixation from the noun <i>međudneвица</i> ‘three-week period between the Assumption and the Nativity of the Blessed Virgin Mary’	In Vuk’s dictionary; <sup>8</sup> in the work of M. Đ. Miličević
2	<i>međunarodan</i> <sup>9</sup> ‘international’	formed after Latin <i>internationalis</i> ‘international’	recent word-formation

8 *Srpski rječnik* (Serbian Dictionary) (1818).

9 The adjective *međunarodan* ‘international’ appears with the *-(a)n* suffix in the Academy’s Dictionary and Benešić’s dictionary, while in the VRH it appears as *međunarodni*, with the *-ni* suffix. An average Croatian speaker would not note any difference in meaning between these two adjectives. An average Croatian speaker with more linguistic knowledge would probably say that *međunarodni* is the definite form of the indefinite adjective *međunarodan*, with no other difference in meaning. Babić (2002: 451–456) has dedicated a whole chapter to the question of differentiating between the *-(a)n* and *-ni* suffixes, which proves in itself that the question is rather complex. These two suffixes present differences on both formal and semantic level. Put briefly, *-(a)n* is used to form descriptive adjectives (e.g. *pametan* ‘intelligent’), while *-ni* is used to form relational adjectives (e.g. *autobusni* ‘pertaining to buses’). Adjectives taking the *-(a)n* suffix have both indefinite and definite forms, can be compared and can have two types of declension (indefinite and definite), while adjectives formed with *-ni* cannot be compared and only have the definite type of declension.

	Adjective	Etymology (as specified by the Academy's Dictionary)	Comment (from the Academy's Dictionary)
3	<i>međusoban</i> : a) 'internal, domestic' b) 'mutual, occurring between people'	-	a) only found in Daničić's dictionary <sup>10</sup> as a quote from a 14 <sup>th</sup> century document b) appears in certain works written by J. Rajić, V. Karadžić, P. Petrović, M. Pavlinović and B. Bogišić
4	<i>međusošan</i> 'interfurcal'	formed after Latin <i>interfurcalis</i> or German <i>gabelständig</i> 'forked'	only in Šulek's DST <sup>11</sup>
5	<i>međustaničan</i> 'intercellular'	formed after Latin <i>intercellularis</i> . e.g. <i>međustanična tvar</i> as equivalent for German <i>Intercellulärsubstanz</i> 'intercellular matter'	only in Šulek's DST
6	<i>međusudan</i> 'interjudicial'	equivalent of Latin <i>interiudicialis</i> 'interjudicial', a Church law term	only found in one writer's work
7	<i>međutiman</i> 'temporary'	<i>među-</i> 'between' + <i>tim</i> 'that'; e.g. <i>Zwischenregierung</i> – <i>međutimna vlada</i> 'interim government'	a recent legal term
8	<i>međuviličan</i> 'intermaxillary'	equivalent of German <i>Zwischenkieferknochen</i> 'intermaxillary bone'	only in Šulek's DST
9	<i>međuzeman</i> 'occurring between countries'	<i>međuzemno more</i> 'sea between countries' as equivalent of German <i>Binnenmeer</i>	only in Šulek's DST
10	<i>međuzemski</i> SYN <i>međuzeman</i> 'intercountry'	equivalent of German <i>Zwischenverkehr</i> 'intercountry traffic'	-

10 *Rječnik iz književnih starina srpskih* (Dictionary of Older Serbian Literature), 1863–1864.

11 *Hrvatsko-njemačko-talijanski rječnik znanstvenoga nazivlja* (Croatian-German-Italian Dictionary of Scientific Terminology) published by Bogoslav Šulek in 1874/1875. The Dictionary often provides French and English, as well as Latin and Greek equivalents. See <http://ihj.hr/iz-povijesti/bogoslav-sulek/38/>.



When one observes the adjectives from Table 1, one immediately understands that, firstly, their number is rather limited, amounting to as few as ten adjectives altogether.<sup>12</sup> Secondly, and more importantly, almost all of the listed Croatian adjectives are claimed to be equivalents of Latin(ate) or German specialized terms, which means these are intentional calques created for the purposes of filling specific lexical gaps. Thirdly, and probably most importantly, five of the total of ten adjectives are hapax legomena, i.e. lexemes found in a single work, whether it be the opus of an uncited writer (*međusudan* ‘interjudicial’) or, for the remaining four, terms coined by Bogoslav Šulek for the purposes of compiling his previously mentioned DST. It needs to be emphasized that Šulek played a large role in the formation of several domains of Croatian scientific terminology, which were a result of real needs for Croatian terms in specific scientific domains and also a way to resist Germanization and/or Hungarization (Samardžija 1997: 178).

A special comment should be made concerning the adjectives *međudnevički* ‘related to *međudnevisa* (‘three-week period between the Assumption and the Nativity of the Blessed Virgin Mary’) and *međusoban* ‘mutual’. The first adjective is derived from the noun *međudnevisa* as a result of suffixation, and thus the prefix *među-* was not directly involved in its formation ( $[međudnevisa]_N + -čki$  ‘suffix’ >  $[međudnevički]_{Adj}$ ). In other words, the prefix *među-* is actually present in the adjective *međudnevički* as part of the noun from which the adjective was derived.

The adjective *međusoban* ‘mutual’ is more interesting. According to the Academy’s Dictionary (1904–1910), at the beginning of the 20<sup>th</sup> century, it had two possible meanings: 1) ‘internal, domestic’ and 2) ‘mutual, occurring between people’. The first meaning is claimed to be found only in Daničić’s dictionary (1863/64) as part of a quote from a 14<sup>th</sup> century Serbian document, while the second appears in some works written by Croatian, Serbian and Montenegrin authors (see footnote 12), who were mostly active during the 19<sup>th</sup> century. The first meaning is not recorded in the Croatian language today,<sup>13</sup> and it was probably never part of it, according to the diachronic information from the Academy’s Dictionary (cf. Matasović et al. 2016: 600). The second meaning of the adjective *međusoban*, ‘mutual’, is the only meaning the adjective has in contemporary Croatian. The adjective was formed according to the [Pref-Pron-Suff]<sub>Adj</sub> pattern, or more precisely according to the following formula: *među-* ‘inter-’ + *sebe* ‘reflexive-possessive pronoun’ + *-(a)n* ‘suffix’ > *međusoban*. It is the only adjective in the Academy’s Dictionary (and in the analysed corpora, as will be said *infra*) that was formed from a pronoun. Thus, the word-formation pattern from which it resulted is an isolated one, and did not have further impact on the formation of Croatian adjectives.

12 One of the anonymous reviewers has asked why there are so few adjectives, and whether the reason lies in their predictability. We do not think that is a plausible answer, firstly, due to the fact that the Academy’s Dictionary is a very comprehensive one, and it would thus list as many possible lexemes as there are; and secondly, because the adjectives from Table 1 mostly belong to specialized languages, which points to the conclusion that *među-* adjectives were rare in general language.

13 Cf. [http://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=e1tjURM%3D&keyword=me%C4%91usoban](http://hjp.znanje.hr/index.php?show=search_by_id&id=e1tjURM%3D&keyword=me%C4%91usoban).

It should also be added that the Academy’s Dictionary lists a number of *među-* prefixed nouns<sup>14</sup> (e.g. *međubrđe* ‘place between hills’ (< *brdo* ‘hill’); *međuvođe* ‘place between waters’ <*voda* ‘water’)), among which many toponyms (e.g. *Međuhan* ‘village in Serbia’; *Međulići* ‘village in Herzegovina’, etc.), thus it seems that in early 20<sup>th</sup> century the *među-* prefix was used only as a noun-deriving prefix.<sup>15</sup>

All of these facts point to the conclusion that, at the beginning of the 20<sup>th</sup> century, a period when the Croatian standard was already formed, there existed very few adjectives formed with the prefix *među-*, and the extant ones were mostly scientific terms related to specific domains of specialized language, as well as hapax legomena. The only two adjectives used more frequently, according to the Academy’s Dictionary, which are not claimed to appear “only once” or “only with a certain author”, are *međusoban* ‘mutual’, which is also the only adjective formed according to the [Pref-Pron-Suff]<sub>Adj</sub> pattern, and *međunarodan* ‘international’, formed as a result of prefix-suffix combination. The latter, however, is “a recent invention” (Budmani/Maretić 1904–1910: 579), which is another interesting insight important for future discussion, as it may lead to the conclusion that it was in the 19<sup>th</sup> and early 20<sup>th</sup> century that *među-* adjectives derived through prefix-suffix combination under foreign influence started to be formed and used.

### 6.1.2 Benešić’s Dictionary

Table 2 presents the relevant adjectives from Benešić’s dictionary (1957).

Table 2: Adjectives formed with *među-* in Benešić’s dictionary (1957).

	Adjective
1	<i>međunarodan</i> ‘international’
2	<i>međusatni</i> ‘between two (school) classes’
3	<i>međuzvjezdani</i> ‘interstellar’

14 A number of these nouns are claimed to be calques of mostly German terms (Budmani/Maretić 1904–1910: 581).

15 A quick search of all words beginning with the sequence *među* in the Klasici subcorpus of HNK, which comprises the greatest classical works of the Croatian literature, gives the following results: *među*<sub>Prep</sub> ‘between’ 1,926 occurrences, *međutim*<sub>Adv</sub> ‘however’ 277 occurrences, *međusoban*<sub>Adj</sub> ‘mutual’ 60 occurrences, *međuto*<sub>Adv</sub> ‘in the meantime’ 7 occurrences, *međunarodan*<sub>Adj</sub> ‘international’ 3 occurrences, *međumurski*<sub>Adj</sub> ‘pertaining to *Međumurje*, the region between the rivers of Drava and Mura’ 2 occurrences, older version of today’s adjective *međimurski* with the same meaning, *međuvrijeme*<sub>N</sub> ‘meantime’ 1 occurrence, *međuakt*<sub>N</sub> ‘entr’act’ 1 occurrence, *međumurec*<sub>N</sub> ‘type of cheese from the region of *Međimurje*’ 1 occurrence. These corpus results show that in older layers of the Croatian language, the prefix *među-* was mostly used to form nouns, except for the two already mentioned adjectives *međunarodan* and *međusoban*, as attested also by the earlier quoted Academy’s Dictionary. The third adjective appearing in the Klasici subcorpus, *međumurski*, is derived through suffixation from the toponym *Međumurje*, so *među-* did not participate directly in its formation.

As Table 2 demonstrates, Benešić’s dictionary (1957) lists only three *među-* adjectives. This fact points to the conclusions that, firstly, over the period from the start of the 20<sup>th</sup> century, i.e. the years when the volume containing *među-* adjectives of the Academy’s dictionary was published, until the mid-20<sup>th</sup> century, when Benešić finished compiling his own dictionary, the pattern [*među-N-Suff*]<sub>Adj</sub> had not yet become a productive one. Secondly, this might have to do with the fact that Benešić’s dictionary is a dictionary of the Croatian literary language, and as such is based on literary sources only, while the observed pattern might already have become more productive in other language areas, such as the press and specialized terminology.

## 6.2 A More Recent Dictionary – VRH (2015)

A more recent dictionary, VRH, lists the following *među-* adjectives.

Table 3: List of *među-* prefixed adjectives in the VRH dictionary (2015: 720–721)

	Adjective
1	<i>međučeljustan</i> ‘intermaxillary’
2	<i>međudisciplinaran</i> ‘interdisciplinary’
3	<i>međudržavni</i> ‘interstate’
4	<i>međufakultetski</i> ‘occurring between faculties’
5	<i>međugalaktički</i> ‘intergalactic’
6	<i>međuglasni</i> ‘intervocalic’
7	<i>međugradski</i> ‘intercity’ (‘inter-city’)
8	<i>međukatni</i> ‘occurring between two or more floors (of a building)’
9	<i>međuljudski</i> ‘occurring between two or more people’
10	<i>međumišićni</i> ‘intermuscular’
11	<i>međumjesni</i> ‘occurring between two or more places’
12	<i>međunarodni</i> ‘international’
13	<i>međuopćinski</i> ‘occurring between two or more municipalities’
14	<i>međuovisan</i> ‘interdependent’
15	<i>međuparlamentaran</i> ‘interparliamentary’
16	<i>međuplanetaran</i> ‘interplanetary’
17	<i>međurasni</i> ‘interracial’
18	<i>međuratni</i> ‘interwar’ (‘inter-war’)
19	<i>međurebreni</i> ‘intercostal’
20	<i>međuregionalni</i> ‘interregional’
21	<i>međurepublički</i> ‘occurring between two or more republics’
22	<i>međuriječni</i> ‘occurring between two or more rivers’

	Adjective
23	<i>međusamoglasnički</i> ‘intervocalic’
24	<i>međusaveznički</i> ‘occurring between two or more allies’
25	<i>međusoban</i> ‘mutual’
26	<i>međuspolan</i> ‘intersexual’
27	<i>međustaničan</i> ‘intercellular’
28	<i>među stranački</i> ‘interparty’ (‘inter-party’)
29	<i>međuvilični</i> ‘intermaxillary’
30	<i>međuvjerski</i> ‘interfaith’ <sup>16</sup>
31	<i>međuvladin</i> ‘intergovernmental’
32	<i>međuvremen</i> ‘occurring between two or more periods of time’
33	<i>međuzavisan</i> ‘interdependent’
34	<i>međuzglobni</i> ‘occurring between joints’
35	<i>međuzubni</i> ‘interdental’
36	<i>međuzvezdan</i> ‘interstellar’
37	<i>međužupanijski</i> ‘occurring between two or more counties’

As Table 3 shows, in the VRH dictionary there are 37 *među-* adjectives, most of which also figure among the ones found in the corpora (see *infra*). However, among them there are also several adjectives absent from our corpus analysis (*međusamoglasnički* ‘intervocalic’, *međumišićni* ‘intermuscular’, *međuriječni* ‘occurring between rivers’, *međuzglobni* ‘occurring between joints’). Upon verification, it was noticed that these four adjectives do not appear in the Riznica corpus at all, and that they appear with a frequency below 10 in the hrWaC corpus. It is therefore interesting that they were chosen to be included in the VRH dictionary while some higher-frequency *među-* adjectives were not.<sup>17</sup>

The 37 adjectives listed in the VRH (2015) demonstrate the following: firstly, 37 adjectives are much more than what can be found in the earlier Academy’s (1904–1910) and Benešić’s dictionary (1957), which testifies itself to the fact that the pattern [*među*-N-Suff]<sub>Adj</sub> has become more productive in Croatian in the recent decades. Secondly, it must also be noted that the dictionary does probably not list all the possible adjectives derived on the basis of that pattern due to the fact that it is a rather predictable and semantically transparent adjective forming process which can virtually use any noun to form a *među-* adjective.

16 One of the anonymous reviewers has enquired about whether the adjective *interfaith* (e.g. <https://www.merriam-webster.com/dictionary/interfaith>) is actually *interreligious* (cf. <https://www.merriam-webster.com/dictionary/interreligious>). In Croatian, *međuvjerski* ‘interfaith’ is derived from the noun *vjera* ‘faith’, which usually refers to Christianity, while *međureligijski* ‘interreligious’ is derived from *religija* ‘religion’, which refers more generally to any belief system. Both English adjectives are listed in Merriam Webster’s dictionary, so we believe they should be distinguished.

17 This is peculiar if one knows that the VRH dictionary claims to be based, among other sources, on three digital corpora: Riznica, HNK and hrWaC (version 2.0).

## 6.3 Corpora Analysis: Riznica, HNK and hrWaC

### 6.3.1 Riznica

In the Riznica corpus, only two adjectives appear in the earliest period covered (between late 19<sup>th</sup> and early 20<sup>th</sup> century): *međunarodan* ‘international’ and *međusoban* ‘mutual’. After that, for a long period going all the way to 1962, these two adjectives remain the only ones formed with the prefix *među-* which appear in the corpus, alongside several *među-* prefixed nouns such as *međuvrijeme* ‘meantime’.

The situation starts to change from the 1970s onwards, when two more adjectives appear: *međunacionalni* ‘occurring between nations’ (1971) and *međuratni* ‘interwar’ (1972). After that, the number of *među-* prefixed adjectives really starts to become much larger from the 2000s, with examples such as the following: *međuljudski* ‘inter-human’, *međeukonfesionalni* ‘interconfessional’, *međugeneracijski* ‘intergenerational’, *međudržavni* ‘interstate’, *međuzvezdani* ‘interstellar’, *međuplanetarni* ‘interplanetary’, *međuledeni* ‘interglacial’, *međustanični* ‘intercellular’, *međuparlamentarni* ‘interparliamentary’, *međurepublički* ‘interrepublican’, *međuvladin* ‘intergovernmental’, *međuetnički* ‘interethnic’, *međustranački* ‘interparty’,<sup>18</sup> etc.

These facts can be said to quite convincingly prove that during the whole 20<sup>th</sup> century *među-*adjectives were rather rare in Croatian, with only *međusoban* ‘mutual’ and *međunarodan* ‘international’ being in use, and that it was only from the 1970s onwards that the pattern started forming a larger number of *među-* adjectives.

This idea leads to the question what reasons there are for the *među-* prefixed adjectives to have become rather frequent in the late 20<sup>th</sup> century. Some answers are provided in the following sections.

### 6.3.2 The HNK Corpus

Table 4: *Među-* adjectives retrieved from HNK

	Adjective	Frequency	Per million frequency
1	<i>međubankarski</i> ‘interbank’ (‘inter-bank’)	91	0.42
2	<i>međubankovni</i> <sup>19</sup> ‘interbank’ (‘inter-bank’)	5	0.02

18 Some of the English equivalents of the Croatian adjectives have a prefix-suffix structure (e.g. *interplanetary*, *interparliamentary*, *intergovernmental*, etc.), while others do not contain any suffix at all (e.g. *interstate*, *intercity*, *interparty* etc.). One of the reasons for this latter fact is that in English, adjectives can be formed through conversion from nouns, which is impossible in Croatian (cf. Babić 2002: 51–52). For a critical review of the so-called *noun-to-adjective conversion*, which some authors only consider as a specific attributive function of nouns, see Balteiro (2007: 45 and further).

19 The two adjectives *međubankarski* and *međubankovni* are synonyms, which differ only with respect to the suffix they are formed with (*-ovni* and *-arski*, respectively), and frequency of use.

	<b>Adjective</b>	<b>Frequency</b>	<b>Per million frequency</b>
3	<i>međučeljusni</i> 'intermaxillary'	13	0.06
4	<i>međudržavni</i> 'interstate'	2,418	11.15
5	<i>međuentitetski</i> 'occurring between two or more entities'	37	0.17
6	<i>međuetažni</i> 'occurring between two or more floors (in a building)'	12	0.06
7	<i>međuetnički</i> 'interethnic'	64	0.30
8	<i>međufazni</i> 'interphase'	10	0.05
9	<i>međugeneracijski</i> 'intergenerational' ( 'inter-generational' )	2,144	9.89
10	<i>međugodišnji</i> 'interannual' ( 'inter-annual' )	17	0.08
11	<i>međugradski</i> 'intercity' ( 'inter-city' )	217	1.00
12	<i>međugraničan</i> 'occurring between borders'	41	0.19
13	<i>međuinstitucionalan</i> 'interinstitutional' ( 'inter-institutional' )	104	0.48
14	<i>međukatni</i> 'occurring between floors (in a building)'	29	0.13
15	<i>međuklupski</i> 'interclub' ( 'inter-club' )	12	0.06
16	<i>međukolodvorski</i> 'occurring between (bus, train etc.) stations'	18	0.08
17	<i>međukontinentalan</i> 'intercontinental'	29	0.13
18	<i>međukorejski</i> 'occurring between the two Koreas'	25	0.12
19	<i>međukrojni</i> 'occurring between two markers'	12	0.06
20	<i>međulaboratorijski</i> 'interlaboratory'	185	0.85

	<b>Adjective</b>	<b>Frequency</b>	<b>Per million frequency</b>
21	<i>međuljudski</i> 'interhuman'	597	2.75
22	<i>međuljuštorni</i> '(of molluscs) intervalvular'	46	0.21
23	<i>međuministarski</i> 'interministerial'	41	0.19
24	<i>međumjesan</i> 'occurring between two or more places'	541	2.50
25	<i>međunacionalan</i> 'occurring between two or more nations'	483	2.23
26	<i>međunarodni</i> 'international'	107,864	497.50
27	<i>međuobalni</i> 'occurring between shores'	10	0.05
28	<i>međuobrtnički</i> 'occurring between craftsmen'	53	0.24
29	<i>međuosovinski</i> 'interaxial'	17	0.08
30	<i>međupalestinski</i> 'inter-Palestinian'	27	0.12
31	<i>međuparlamentaran</i> 'interparliamentary'	137	0.63
32	<i>međupopisni</i> 'intercensal'	19	0.09
33	<i>međuratni</i> 'interwar' ('interwar')	142	0.65
34	<i>međuregionalni</i> 'interregional' ('inter-regional')	90	0.42
35	<i>međureligijski</i> 'interreligious' ('inter-religious')	204	0.94
36	<i>međurepublički</i> 'occurring between two or more republics'	45	0.21
37	<i>međuresorni</i> 'occurring between two or more departments'	85	0.39
38	<i>međuresorski</i> 'occurring between two or more departments'	319	1.47
39	<i>međusatelitski</i> 'occurring between two or more satellites'	11	0.05

	<b>Adjective</b>	<b>Frequency</b>	<b>Per million frequency</b>
40	<i>međusektorski</i> 'occurring between two or more sectors'	12	0.06
41	<i>međusoban</i> 'mutual'	20,277	93.52
42	<i>međuspojni</i> 'occurring between links'	12	0.06
43	<i>među stranački</i> 'interparty' ('inter-party')	510	2.35
44	<i>međususjedski</i> 'occurring between neighbours'	50	0.23
45	<i>međuvalutni</i> 'occurring between two or more currencies'	15	0.07
46	<i>međuvjerski</i> 'interfaith'	91	0.42
47	<i>međuvladin</i> 'intergovernmental'	65	0.30
48	<i>međuvlasnički</i> 'occurring between two or more owners'	22	0.10
49	<i>međuzrnski</i> 'intergranular'	19	0.09
50	<i>međužupanijski</i> 'occurring between two or more counties'	332	1.53
51	<i>međuzvezdan</i> 'interstellar'	29	0.13

As Table 4 demonstrates, in the HNK corpus there are 51 *među-* adjectives with a frequency of 10 or more occurrences. That is quite a large number of adjectives formed mostly according to the [*među-N-Suff*]<sub>Adj</sub> pattern, which proves that the pattern became more productive starting from the 1990s. While some of the adjectives from Table 4 belong to specialized languages (e.g. *međuljuštarni* '(of molluscs) intervalvular'; *međuzrnski* 'intergranular'; *međuzvezdani* 'interstellar', etc.), a large number belong to general language. If these adjectives have been found in the national corpus, that points to the conclusion that adjectives formed according to the [*među-N-Suff*]<sub>Adj</sub> pattern have been quite widespread in various genres of standard Croatian since the 1990s.

Finally, the hrWaC corpus shall be analysed as the largest extant corpus of the Croatian language. It will be interesting to compare the *među-* adjectives appearing in hrWaC, as a large web corpus, with those appearing in the national corpus, due to the different text types they include.



### 6.3.3 hrWaC

Considering the fact that hrWaC is the largest corpus that registers the most recent use of the Croatian language, its inclusion in the analysis presented in this paper was important for learning about the current use of adjectives formed with *među-*. What the corpus data shows is that 134 adjectives with the prefix *među-* and a frequency of 10+ occurrences are registered in hrWaC. 134 is a lot more than only two adjectives in the first half of the 20<sup>th</sup> century, and much more than just several more around the 1970s. This means that the derivational pattern [*među-N-Suff*]<sub>Adj</sub> has become more productive with time, culminating in recent years.

Below is a list of the 50 most frequent adjectives retrieved from hrWaC.

Table 5: The 50 most frequent *među-* adjectives retrieved from hrWaC.

	<b>Adjective</b>	<b>Frequency</b>	<b>Per million frequency</b>
1	<i>međunarodni</i> ‘international’	222,469	214.80
2	<i>međusoban</i> ‘mutual’	48,796	36.62
3	<i>međuljudski</i> ‘interhuman’	6,837	5.15
4	<i>međudržavni</i> ‘interstate’	4,539	3.25
5	<i>međugeneracijski</i> ‘intergenerational’	4,384	3.14
6	<i>međuosovinski</i> ‘interaxial’	2,716	1.94
7	<i>međužupanijski</i> ‘occurring between two or more counties’	2,076	1.49
8	<i>međunacionalni</i> ‘occurring between two or more nations’	1,747	1.25
9	<i>međuvladin</i> ‘intergovernmental’	1,508	1.08
10	<i>međureligijski</i> ‘interreligious’	1,235	0.88
11	<i>međustranački</i> ‘interparty’	1,030	0.74
12	<i>međugradski</i> ‘intercity’	1,019	0.73
13	<i>međuetnički</i> ‘interethnic’	697	0.50
14	<i>međuratni</i> ‘interwar’	687	0.49
15	<i>međusektorski</i> ‘occurring between two or more sectors’	671	0.48
16	<i>međuzvjezdani</i> ‘interstellar’	603	0.43
17	<i>međubankarski</i> ‘interbank’	576	0.41
18	<i>međuvlasnički</i> ‘occurring between two or more owners’	556	0.40
19	<i>međuregionalan</i> ‘interregional’	555	0.40

	<b>Adjective</b>	<b>Frequency</b>	<b>Per million frequency</b>
20	<i>međuopćinski</i> ‘occurring between two or more municipalities’	520	0.37
21	<i>međuresorni</i> ‘occurring between two or more departments’	483	0.35
22	<i>međukulturalni</i> ‘intercultural’	457	0.33
23	<i>međukatni</i> ‘occurring between two or more floors (in a building)’	406	0.29
24	<i>međuvjerski</i> ‘interfaith’	391	0.28
25	<i>međustanični</i> ‘intercellular’	374	0.27
26	<i>međugodišnji</i> ‘interannual’	348	0.25
27	<i>međumjesni</i> ‘occurring between two or more communities’	318	0.23
28	<i>međukulturni</i> ‘intercultural’	297	0.21
29	<i>međuplanetarni</i> ‘interplanetary’	295	0.21
30	<i>međuknjižnični</i> ‘occurring between two or more libraries’	295	0.21
31	<i>međuovisan</i> ‘mutually dependent’	241	0.17
32	<i>međuinstitucionalan</i> ‘interinstitutional’	222	0.16
33	<i>međurepublički</i> ‘occurring between two or more republics’	217	0.16
34	<i>međugranični</i> ‘occurring between two or more borders’	215	0.15
35	<i>međuresorski</i> ‘occurring between two or more departments’	208	0.15
36	<i>međususjedski</i> ‘occurring between two or more neighbours’	198	0.14
37	<i>međurasni</i> ‘interracial’	198	0.14
38	<i>međubankovni</i> ‘interbank’	195	0.14
39	<i>međuparlamentaran</i> ‘interparliamentary’	193	0.14
40	<i>međuškolski</i> ‘occurring between two or more schools’	182	0.13
41	<i>međuzubni</i> ‘interdental’	150	0.11
42	<i>međulaboratorijski</i> ‘occurring between two or more laboratories’	159	0.11
43	<i>međuvršnjački</i> ‘occurring between peers’	134	0.10
44	<i>međugalaktički</i> ‘intergalactic’	133	0.10

	Adjective	Frequency	Per million frequency
45	<i>međuredni</i> ‘occurring between two or more lines’	131	0.09
46	<i>međupredmetni</i> ‘occurring between two or more subjects’	123	0.09
47	<i>međunožni</i> ‘occurring between legs’	115	0.08
48	<i>međuentitetski</i> ‘occurring between two or more entities’	113	0.08
49	<i>međuklupski</i> ‘interclub’ (‘inter-club’)	107	0.08
50	<i>međurebreni</i> ‘intercostal’	94	0.07

The 134 *među-* adjectives extracted from the hrWaC corpus is by far the largest number of adjectives appearing in any of the aforementioned sources. Such an outcome was expected, firstly, because hrWaC is the largest extant Croatian corpus. Secondly, given the fact that hrWaC is a web corpus, which also collects data from specific sources such as chat rooms and blogs, we have expected to find a certain portion of less standard or more colloquial terms, e.g. *međunožni* ‘occurring between legs’, but also more specific terms such as *međuroditeljski* ‘interparental’, which seems to appear in a single weekly newspaper; or *međuizborni* ‘occurring between elections’, which is used exclusively in political discourse. It should also be emphasized that some of the adjectives retrieved from hrWaC belong to highly specialized terminologies, such as *međuzrnski* ‘intergranular’, *međubiskupijski* ‘interdiocesan’ or *međukralježnički* ‘inter-spinal’, which one would not expect to find in general language dictionaries such as the VRH, but rather in specialized dictionaries or glossaries.

It can generally be said that the 134 adjectives retrieved from hrWaC point to the conclusion that *među-* adjectives, a majority of which were formed according to the [*među-N-Suff*]<sub>Adj</sub> pattern, have recently become quite numerous in Croatian. The fact also implies that the [*među-N-Suff*]<sub>Adj</sub> pattern continues to produce new forms, some of which have not been listed in lexicographic works yet. It would be interesting to analyse *među-* adjectives below the 10 occurrences threshold to see what happens at that end of the frequency scale, how many adjectives there are and what characteristics they have. It can be assumed that there would be more new adjectives that have not been listed in dictionaries yet due to the fact that the pattern seems to be quite productive.

The adjectives extracted from hrWaC shall now be analysed first morphologically, or from a purely word-formation point of view, and then semantically.

#### 6.4 Morphological Analysis

From a morphological point of view, the adjectives retrieved from hrWaC are a result of:

- 1) prefix-suffix formation according to the formulas [*među-N-Suff*]<sub>Adj</sub> and [*među-Pron-Suff*]<sub>Adj</sub>
- 2) prefixation according to the formula [*među-Adj*]<sub>Adj</sub>.

A great majority of adjectives belong to the first group. They are a result of the simultaneous addition of the prefix *među-* and a suffix to a nominal base (cf. Babić 2002: 445). Thirteen suffixes participated in their formation, as listed in the table below.

Table 6: Suffixes involved in the prefix-suffix formation of adjectives retrieved from hrWaC.

	Suffix	Adjective example	No.	%
1	<i>-ni</i>	<i>međunarodni</i> ‘international’	49	36.6
2	<i>-ski</i>	<i>međugeneracijski</i> ‘intergenerational’	48	35.8
3	<i>-čki</i>	<i>međustranački</i> ‘interparty’	10	7.5
4	<i>-en</i>	<i>međuvremen</i> ‘occurring between two or more periods of time’	6	4.5
5	<i>-ionalan</i>	<i>međunacionalan</i> ‘occurring between two or more nations’	6	4.5
6	<i>-aran</i>	<i>međuparlamentaran</i> ‘interparliamentary’	3	2.2
7	<i>-(a)n</i>	<i>međuzvezdan</i> ‘interstellar’	3	2.2
8	<i>-ovni</i>	<i>međubankovni</i> ‘interbank’	2	1.5
9	<i>-enski</i>	<i>međuplemenski</i> ‘intertribal’	2	1.5
10	<i>-alan</i>	<i>međukontinentalan</i> ‘intercontinental’	2	1.5
11	<i>-šnji</i>	<i>međugodišnji</i> ‘interannual’	1	0.7
12	<i>-in</i>	<i>međuvladin</i> ‘intergovernmental’	1	0.7
13	<i>-ovski</i>	<i>međužanrovski</i> ‘occurring between two or more genres’	1	0.7
<b>Total</b>			<b>134</b>	<b>100%</b>

Table 6 demonstrates that the two most frequent suffixes, which account for the formation of more than 72% percent of all adjectives, are *-ni* and *-ski*. This fact is consonant with details related to the prefix-suffix formation of relational adjectives provided by Babić (2002: 444). Babić (2002: 398) describes the *-ski* suffix as “one of the most productive adjectival suffixes”, and *-ni* as “very productive” (2002: 421). With respect to the use of the *-ski* and *-ni* suffixes in the formation of adjectives, which is a highly complex question dependent on both formal (phonological) and semantic criteria, the following can be said briefly (cf. Babić 2002: 428–434):

- 1) all adjectives derived from proper nouns are formed with *-ski*, never with *-ni*;
- 2) adjectives derived from common nouns related to living beings (humans, animals and plants) are rarely formed with *-ni*;
- 3) both *-ni* and *-ski* are used to form adjectives from common nouns related to non-living entities (both concrete and abstract), and their distribution depends mostly on phonological context,<sup>20</sup>

20 For instance, adjectives from nouns ending in *-ij* are formed with *-ski*: e.g. *natrij* ‘natrium’ > *natrijski* ‘related to natrium’; *laboratorij* ‘laboratory’ > *laboratorijski* ‘related to laboratory’, etc. (Babić 2002: 429). The *-ni* suffix is used, for example, with bases ending in *-st* or *-št*: e.g. *čeljust*

- 4) a number of adjectives have dual forms with *-ni* and *-ski*, but with different frequency of use, some of which are also non-standard.

The only adjective formed according to the  $[me\dot{d}u\text{-Pron-Suff}]_{Adj}$  pattern, i.e. from a pronominal base, is *međusoban* ‘mutual’, which is also the only descriptive adjective formed with *među-* through prefix-suffix combination, according to Babić (2002: 461).

With respect to the second group of adjectives, formed through prefixation according to the formula  $[me\dot{d}u\text{-Adj}]_{Adj}$ , only the following three adjectives exemplify it: *međuzavisan* ‘interdependent’, *međuovisan* ‘interdependent’<sup>21</sup> and *međupovezan* ‘interconnected’. They are formed via the addition of the prefix *među-* to a base without the participation of a suffix. These three adjectives account for only 2.2% of all the *među-*prefixed adjectives in the corpus.

### 6.5 Semantic Analysis

Without entering into details, one can observe that the prefix *među-* connects with nominal bases from various semantic domains to form adjectives, resulting in terms related to administration (e.g. *međuopćinski* ‘occurring between municipalities’), traffic (e.g. *međukontinentalan* ‘intercontinental’), zoology (e.g. *međutelidbeni* ‘(of cows) intercalving’), finance (e.g. *međuvalutni* ‘intercurrency’), astronomy (e.g. *međugalaktički* ‘intergalactic’), politics (e.g. *međustranački* ‘interparty’), religion (*međureligijski* ‘interreligious’), etc.

What interests us more here is prefixal meanings. The prefix *među-* realizes two types of meanings in the analysed adjectives: concrete and abstract. Its concrete meaning is ‘located between two or more (concrete) entities’ (e.g. *međustaklen* ‘between two glass surfaces’). This is the core or prototype (e.g. Lakoff 1987) meaning that refers to the concrete spatial position of concrete objects.

The abstract meanings of the prefix *među-* in the analysed adjectives are the following:

- a) ‘between two or more abstract entities’ (e.g. *međugeneracijska solidarnost* ‘intergenerational solidarity’), and
- b) ‘between two or more periods of time’ (e.g. *međutelidbeno razdoblje* ‘intercalving period’).

The semantic network that the prefix *među-* realizes with adjectives can thus be illustrated by the following image.

---

‘jaw’ > *čeljusni* ‘related to jaw’; *kazalište* ‘theatre’ > *kazališni* ‘related to theatre’, etc. (Babić 2002: 430).

21 The adjectives *međuzavisan* and *međuovisan* are near-synonyms and can be used interchangeably in most contexts. Cf. [http://hjp.znanje.hr/index.php?show=search\\_by\\_id&id=e1tjUBQ%3D&&keyword=me%C4%91uzavisan](http://hjp.znanje.hr/index.php?show=search_by_id&id=e1tjUBQ%3D&&keyword=me%C4%91uzavisan).

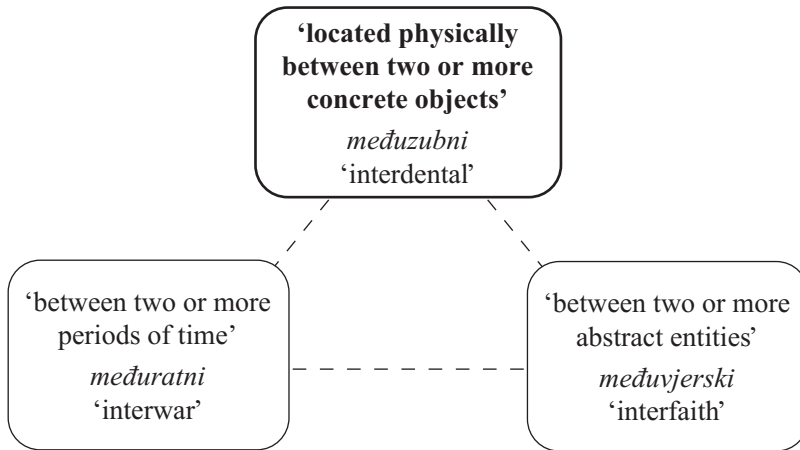


Figure 1: Semantic network of the prefix *među-* in the analysed adjectives.

The abstract meaning ‘between two abstract entities’ relies upon the abstract is concrete metaphor (e.g. Lakoff 1987) on the basis of which we conceptualize more abstract entities through more concrete ones. In examples such as *međuvjerski dijalog* ‘interfaith dialogue’ or *međukulturno razumijevanje* ‘intercultural understanding’, for instance, we perceive abstract phenomena such as dialogue and understanding taking place between faiths or cultures as phenomena occurring between two concrete things, thus we express them with the same preposition (*među* ‘between’) or prepositional prefix *među-* ‘inter-’.

The concrete meaning ‘located physically between two or more concrete objects’ is metaphorically extended into the meaning ‘between two periods in time’ on the basis of the frequent time is space metaphor (e.g. Kövecses 2010). This metaphor enables human beings to conceptualize time phenomena on the basis of concrete, spatial phenomena of which they have better understanding. Thus, in examples from hrWaC such as *međusezonska kolekcija* ‘interseasonal (clothing or shoes) collection’ what happens between two periods of time, i.e. two seasons, is conceptualized as being physically located between two concrete objects.

In the conclusion to this part, it must be emphasized that the prefix *među-* is a polysemous affix which, when attached to adjectives, realizes three related meanings, both concrete and abstract, the latter of which are based on metaphor. The semantic network of the prefix *među-* in the analysed adjectives demonstrates that it behaves much like other lexical categories such as nouns and verbs, construing a radial structure with the prototypical sense as the centre of its semantic network (cf. Tyler and Evans 2003: 31). The semantic level of the formation of the analysed adjectives was insisted upon owing to the fact that, according to our understanding, all complex words are motivated both grammatically (or morphologically) and semantically, i.e. that derivational processes cannot be separated from the semantic ones (cf. Booij 2005; Raffaelli 2013).<sup>22</sup>

<sup>22</sup> For a different view on word-formation, see Aronoff (1976) and Scalise (1984) among others.

The semantic network of this particular prefix is just an example of the complex relations that exist between word-formation processes and meanings that are created during the derivation of new complex lexemes. Therefore, some authors (e.g. Raffaelli 2018: 153) emphasize that one of the major future tasks of word-formation as a linguistic subdiscipline is to systematically study the semantic processes which accompany the formation of complex words.

## 7 CONCLUDING REMARKS

This paper explores Croatian adjectives formed with the prefix *među-*. While adjectives derived according to the  $[među-N-Suff]_{Adj}$  pattern are fairly numerous in contemporary Croatian, according to Babić (2002), the author of the most comprehensive manual of Croatian word-formation, as well as lexicographic sources (the Academy's Dictionary, Benešić's dictionary and the VRH) and large digital corpora (hrWaC and HNK), from a diachronic point of view that was not the case as recently as only a hundred years ago. More precisely, both an analysis of older lexicographic works (the Academy's Dictionary and Benešić's dictionary), as well as of digital corpora covering older texts (Riznica and HNK) have showed that in early 20<sup>th</sup> century *među-* prefixed adjectives were very rare. Moreover, the Academy's Dictionary (1904–1910) specifies that almost all such adjectives were hapaxes and calques made according to Latin(ate) or German models. In mid-20<sup>th</sup> century, the situation was rather similar according to both dictionaries and corpora, and adjectives formed with *među-* only seem to have become more numerous later in the 20<sup>th</sup> century. The question is, obviously, why.

The Academy's Dictionary's explicit claim that all adjectives formed with *među-* but one, *međusoban* 'mutual', are hapaxes and equivalents of foreign terms is a very important one, because it points to a temporarily conclusion that these are not native Croatian formations. It must be added immediately, however, that there were a number of nouns in the same period that were formed with *među-*. In other words, it seems that *među-* used to be an exclusively noun-forming prefix, which was impossible to be used with adjectives before the 19<sup>th</sup> century. When it comes to the adjective *međusoban* 'mutual', it was demonstrated that it was used with two meanings, the first of which, 'domestic', appears in a single 14<sup>th</sup> century Serbian document, and the second of which, 'mutual', which is the meaning it still has in contemporary Croatian, has been registered in texts of Croatian, Montenegrin and Serbian authors mostly from the 19<sup>th</sup> century. In other words, even the adjective *međusoban* 'mutual', as an isolated derivation resulting from a prefix-suffix combination with a pronominal base (i.e. from the reflexive-possessive pronoun *se(be)*), is a creation dating back to the 19<sup>th</sup> century, as the rest of *među-* prefixed adjectives.

Apart from the case of the adjective *međusoban* 'mutual', the influence of foreign languages seems to be a key element contributing to the possibility of adjective formation with the prefix *među-*. While Latin *inter-* 'between' was the earliest language source of Croatian *među-* prefixed adjective calques, followed by few German words formed with *zwischen-* 'inter-', neither of these languages seemed to provide a large number of new adjectives in Croatian. It was only in late 20<sup>th</sup> century, which coincided with the advent

of English predominance and the spread of its influence on Croatian (Filipović 1990; Muhvić Dimanovski 1992: 94; Samardžija 2002: 72), that the Croatian language saw a large number of new *među-* prefixed adjectives. It is therefore arguably under English influence that numerous adjectives with *među-* were formed and then spread in Croatian from mid-20<sup>th</sup> century onwards, reinforcing the derivational pattern [*među-N-Suff*]<sub>Adj</sub> which had already existed in the language as a result of early Latin(ate) and German calques. The Etymological Dictionary of the Croatian Language (Matasović et al. 2016)<sup>23</sup> also hints at this. Moreover, Ranko Matasović (p. c.),<sup>24</sup> one of the leading experts on the history of the Croatian language and Slavonic languages in general, considers that the pattern has recently become productive in Croatian, and that several decades ago most of *među-* adjectives could have been formed as calques of English adjectives.

A subsequent question would be whether this particular derivational pattern came to be through indirect or direct English influence (cf. Seifart 2015). While the question cannot be answered with certainty, it was probably a case of direct borrowing, or direct calque, as English has been a rather widespread language among Croatian speakers, and the most spoken foreign one, in the last decades. More precisely, due to increasingly intensive contacts with English *inter-* prefixed adjectives, Croatian speakers have probably started to calque them in the domains they needed them to fill in lexical gaps, using the Croatian prefix *među-* coupled with Croatian nominal bases and a suffix.

At some point in time, the derivational pattern [*među-N-Suff*]<sub>Adj</sub> could probably have become as “natural” as any other adjective-deriving native Croatian pattern. If this scenario were correct, it would not be a case of the introduction of a new element in the Croatian language, but of a reinforcement of an existing prefix (*među-*) in a new “surrounding”, i.e. with adjectival bases, because the prefix had been used to form complex nouns in the 19<sup>th</sup> century and even earlier (for instance in a number of toponyms).

The analysis of complex *među-* prefixed adjectives presented in this paper demonstrates that the adjective-forming pattern [*među-N-Suff*]<sub>Adj</sub>, which has entered the Croatian language as a result of loan translation of Latin(ate) and German terms, and was subsequently probably reinforced through the calquing of English *inter-* adjectives, is a productive word-formation pattern in contemporary Croatian. More precisely, today it accounts for a number of adjectives belonging to semantically various domains, according to Croatian word-formation manuals and recent lexicographic works. Moreover, it continues to produce new adjectives, as attested by large web corpora. Not only do these insights illustrate the complex influences foreign languages (such as Latin, German and English) have had on the Croatian word-formation and lexicon, but they also make a contribution, however modest, to the study of morphological borrowing as a phenomenon in general.

23 The dictionary lists only three *među-* adjectives: *međunarodni* ‘international’, which it claims to be a calque of English and French *international*, *međusoban* ‘mutual’, derived from *među* ‘between’ and *se* ‘oneself’, and *međugradski* ‘inter-city’, for which it claims that it was formed after the English adjective *inter-city* (Matasović et al. 2016: 600).

24 I would hereby like to thank Ranko Matasović, fellow of the Croatian Academy of Sciences and Arts, for having kindly shared his helpful insights regarding the question addressed in this paper.



## Primary sources

Hrvatski jezični portal (Croatian Language Portal). August 2020. <http://hjp.znanje.hr/>  
HNK (Croatian National Corpus). August 2020. [http://filip.ffzg.hr/cgi-bin/run.cgi/first\\_form?corpname=HNK\\_v30;align=](http://filip.ffzg.hr/cgi-bin/run.cgi/first_form?corpname=HNK_v30;align=)  
hrWaC Croatian web corpus. August 2020. <http://nlp.ffzg.hr/resources/corpora/hrwac/>  
Riznica corpus. August 2020. <http://riznica.ihjj.hr/>

## References

- ARONOFF, Mark (1976) *Word Formation in Generative Grammar*. Massachusetts: Massachusetts Institute of Technology.
- BABIĆ, Stjepan (2002) *Tvorba riječi u hrvatskome književnome jeziku*. 2. izdanje. Zagreb: Nakladni zavod Globus/Hrvatska akademija znanosti i umjetnosti.
- BALTEIRO, Isabel (2007) *The Directionality of Conversion in English: A Diachronic Study*. Bern: Peter Lang.
- BOOIJ, Geert E. (2005) *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford/New York: Oxford University Press.
- BROZOVIĆ RONČEVIĆ, Dunja/Damir ČAVAR (2012) “Riznica: The Croatian Language Corpus.” *Prace filologiczne* 63, 51–65.
- BUDMANI, Pero/Tomo MARETIĆ (1904–1910) *Rječnik hrvatskoga ili srpskog jezika. Dio VI. Lekenički – Moračice*. Zagreb: Jugoslavenska akademija znanosti i umjetnosti.
- COGHILL, Eleanor (2015) “Borrowing of verbal derivational morphology between Semitic languages: the case of Arabic verb derivations in Neo-Aramaic.” In: F. Gardani/P. Arkadiev/N. Amiridze (eds), 83–108.
- DABO-DENEGRI, Ljuba (2007) *Hrvatsko-francuski jezični dodiri: s rječnikom galicizama u hrvatskom standardnom jeziku*. Zagreb: Nakladni zavod Globus.
- DELLA VOLPE, Angela (1997) “Should we ‘re-consider’? A look at the pragmatics of the semantics of re-.” In: R. Hickey/S. Puppel (eds), *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on his 60th birthday. Volume II: Linguistic Modelling*. Berlin/New York: Mouton de Gruyter, 1653–1665.
- FILIPOVIĆ, Rudolf (1990) *Anglicizmi u hrvatskom ili srpskom jeziku*. Zagreb: Jugoslavenska akademija znanosti i umjetnosti/Školska knjiga.
- GARDANI, Francesco/Peter ARKADIEV/Nino AMIRIDZE (2015) “Borrowed morphology: an overview.” In: F. Gardani/P. Arkadiev/N. Amiridze, 1–24.
- GOSTL, Igor (1995) *Bogoslav Šulek: otac hrvatskoga znanstvenoga nazivlja*. Zagreb: Matica hrvatska.
- GRANT, Anthony P. (2019) “Contact-Induced Linguistic Change: An Introduction.” In: A. P. Grant (ed), *The Oxford Handbook on Language Contact*. New York: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199945092.013.1>
- IACOBINI, Claudio (2020) “Parasynthesis in morphology.” In: *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press, 1–20. <https://doi.org/10.1093/acrefore/9780199384655.013.509>

- JAKOBSON, Roman (1938) "Sur la théorie des affinités phonologiques entre les langues." In: *Actes du quatrième Congrès International de Linguistes, tenu à Copenhague du 27 août au 1er septembre 1936*. Copenhagen: Einar Munksgaard, 48–59.
- KOVAČEC, August (1967) "Bilingvizam i aloglotski utjecaji na morfosintaktičku strukturu jezika (na istrorumunjskom materijalu)." *Suvremena lingvistika* 4, 101–114.
- KÖVECSES, Zoltan (2010) "Metaphor, language, and culture." *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada* 26, 739–757. <https://doi.org/10.1590/S0102-44502010000300017>
- LAKOFF, George (1987) *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*. Chicago/London: The University of Chicago Press.
- LJUBEŠIĆ, Nikola/ERJAVEC, Tomaž (2011) "hrWaC and slWaC: Compiling Web Corpora for Croatian and Slovene." In: I. Habernal/V. Matousek (eds), *Speech and Dialogue, Lecture Notes in Computer Science*. Berlin/Heidelberg: Springer, 395–402.
- MARGIĆ DRLJAČA, Branka (2009) "Latentno posuđivanje u hrvatskome i drugim jezicima – posljedice i otpori." *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 35/1, 53–71. <https://hrcak.srce.hr/50700>
- MARTINET, André (1980) *Eléments de linguistique générale*. Paris: Armand Colin.
- MATASOVIĆ, Ranko/Tijmen PRONK/Dubravka IVŠIĆ/Dunja BROZOVIĆ RONČEVIĆ (2016) *Etimološki rječnik hrvatskoga jezika. I. svezak. A–Nj*. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- MEILLET, Antoine (1926) *Linguistique historique et linguistique générale*. Paris: H. Champion.
- MINAYEVA, Vera (2003) "Russian grammatical interference in Ket." *STUF - Language Typology and Universals* 56/1-2, 40–54. <https://doi.org/10.1524/stuf.2003.56.12.40>
- MUHVIĆ-DIMANOVSKI, Vesna (1992) "Prevedenice – jedan oblik neologizama." *Rad HAZU* 446, 93–205. <https://dizbi.hazu.hr/a/?pr=iiif.v.a&id=18809>
- NIKOLIĆ-HOYT, Anja (2010) "Uz dovršavanje Benešičeva rječnika." *Filologija* 55, 61–87. <https://hrcak.srce.hr/65800>
- POPLACK, Shana (2018) *Borrowing. Loanwords in the speech community and in the grammar*. Oxford/New York: Oxford University Press.
- RAFFAELLI, Ida (2013) "The model of morphosemantic patterns in the description of lexical architecture." *Lingue e linguaggio* 1, 47–72.
- RAFFAELLI, Ida (2018) "Kroz leksikologiju i semantiku". In: P. Košutar/M. Kovačić (eds), *Od dvojbe do razdvojbe. Zbornik radova u čast profesorici Branki Tafri*. Zagreb: Ibis grafika, 143–159.
- SAKEL, Jeannette (2007) "Types of loan: matter and pattern." In: Y. Matras/J. Sakel (eds), *Grammatical Borrowing in Cross-Linguistic Perspective*. Berlin/New York: Mouton de Gruyter, 15–30.
- SAMARDŽIJA, Marko (1997) "Utjecaj sociopolitičkih mijena na leksik hrvatskoga jezika u XX. stoljeću." *Croatica* 45–46, 177–192. <https://hrcak.srce.hr/214731>
- SAMARDŽIJA, Marko (2002) *Nekoć i nedavno. Odabrane teme iz leksikologije i novije povijesti hrvatskoga standardnoga jezika*. Rijeka: Izdavački centar Rijeka.

- SCALISE, Sergio (1984) *Generative Morphology*. Dordrecht: De Gruyter.
- SEIFART, Frank (2015) “Direct and indirect affix borrowing.” *Language* 91/3, 511–532. <https://doi.org/10.1353/LAN.2015.0044>
- SERRANO-DOLADER, David (2015) “Parasynthesis in Romance.” In: P. O. Müller/I. Ohnheiser/S. Olsen/F. Rainer. *Word formation: An international handbook of the languages of Europe*. Berlin: De Gruyter, 524–536.
- SLIŠKOVIĆ, Andrea (2016) “Hrvatski leksikografski vrhovi. Veliki rječnik hrvatskoga standardnog jezika.” *Fluminensia* 28/1, 244–275. <https://hrcak.srce.hr/161103>
- TADIĆ, Marko (2002) “Building the Croatian National Corpus.” In: M. González Rodríguez/C. P. Suarez Araujo (eds), *Third International Conference on Language Resources and Evaluation LREC2002*, Paris/Las Palmas: ELRA, 441–446.
- TADIĆ, Marko (2009) “New version of the Croatian National Corpus.” In: D. Hlaváčková/A. Horák/K. Osolsobě/P. Rychlý (eds), *After Half a Century of Slavonic Natural Language Processing*. Brno: Masaryk University, 199–205.
- THOMASON, Sarah Grey (2001) *Language contact: An introduction*. Washington, DC: Georgetown University Press.
- TRAJDER, Wojciech (2007) “Procédé du calque des nouveaux dérivés préfixaux français en ‘euro-langage’ polonais.” *Studia Romanica Posnaniensia* 34, 137–150. <http://pressto.amu.edu.pl/index.php/srp/article/download/9799/9426>
- TURK, Marija (2013) *Jezično kalkiranje u teoriji i praksi. Prilog lingvistici jezičnih dodira*. Zagreb: Hrvatska sveučilišna naklada/Filozofski fakultet u Rijeci.
- TYLER, Andrea/Vyvyan EVANS (2003) *The Semantics of English Prepositions. Spatial Scenes, Embodied Meaning and Cognition*. Edinburgh: Edinburgh University Press.
- WEINREICH, Uriel (1963). *Languages in Contact: Findings and Problems*. New York: Linguistic Circle of New York.
- ZOMBORI, Lajos/Hernik STEINMANN (1999) *Handbook of Zoology. Volume IV: Anthropoda/Insecta*. Berlin/New York: Walter De Gruyter.

#### Abstract

#### DEVELOPMENT OF A PRODUCTIVE DERIVATIONAL PATTERN ON THE BASIS OF LOAN TRANSLATION? THE CASE OF CROATIAN ADJECTIVES FORMED WITH THE PREFIX *MEDU-*

This paper deals with the question of the formation of Croatian adjectives with the prefix *među-*. While such adjectives were very rare in late 19<sup>th</sup> and early 20<sup>th</sup> century, an analysis of relevant lexicographic works and digital corpora demonstrated that their number started to become larger in later 20<sup>th</sup> century, culminating in recent decades. Today, the [*među-N-Suff*]<sub>Adj</sub> derivational pattern is a productive, accounting for 134 adjectives with a frequency of ten occurrences or more retrieved from the largest extant Croatian web corpus, hrWaC. On the basis of an analysis of available older lexicographic works and digital corpora, it can be concluded that *među-* prefixed adjectives

first entered into Croatian as loan translations (calques) of Latin(ate) and German terms. According to more recent lexicographic works and digital corpora, later on, and especially in recent decades, which coincided with a growing English influence on Croatian, *među-* prefixed adjectives were probably produced as equivalents of English *inter-* prefixed adjectives. The number of *među-* prefixed adjectives, as well as the variety of semantic domains in which they are used, testify to the fact that the [*među-N-Suff*]<sub>Adj</sub> pattern is well-established and productive in contemporary Croatian. The analysis of Croatian *među-* prefixed adjectives in this paper could contribute to shedding more light on the question of morphological borrowing phenomena in general.

**Keywords:** derivational pattern, adjective formation, loan translation (calque), Croatian, language contact

#### Povzetek

### RAZVOJ PRODUKTIVNEGA DERIVACIJSKEGA VZORCA NA PODLAGI IZPOSOJENK? PRIMER HRVAŠKIH PRIDEVNIKOV, IZPELJANIH S PREDPONO *MEDU-*

Prispevek obravnava izpeljavo hrvaških pridevnikov s predpono *među-*. Medtem ko so bili takšni pridevniki v 19. in 20. stoletju redki, razčlemba sodobnih leksikografskih virov in digitalnih korpusov pokaže, da se je njihovo število začelo povečevati v poznem 20. stoletju, sploh pa v zadnjih desetletjih. Danes je torej derivacijski vzorec [*među-N-Suff*]<sub>Adj</sub> v hrvaščini produktiven, saj v trenutno največjem hrvaškem spletnem korpusu hrWaC najdemo 134 takšnih pridevnikov s pogostnostjo nad 10. Razčlemba starejših leksikografskih virov in digitalnih korpusov pokaže, da so se pridevniki s predpono *među-* v hrvaščini najprej pojavili kot izposojenke (kalki) latinskih in nemških izrazov. V novejših leksikografskih virih in digitalnih korpusih pa so se kasneje, sploh v zadnjih desetletjih, ko se povečuje vpliv angleščine na hrvaščino, pridevniki z *među-* verjetno pojavili kot ustreznice angleških pridevnikov s predpono *inter-*. Število pridevnikov s predpono *među-* in različna pomenska polja, kjer se uporabljajo, pričajo o dejstvu, da je vzorec [*među-N-Suff*]<sub>Adj</sub> v sodobni hrvaščini ustaljen in produktiven. Pričujoča analiza hrvaških pridevnikov s predpono *među-* prispeva tudi k razumevanju morfološkega izposojanja na splošno.

**Ključne besede:** derivacijski vzorec, izpeljava pridevnikov, izposojenke (kalki), hrvaščina, jezikovni stik

**Gergana Popova\***  
Goldsmiths,  
University of London, UK

UDK 811.163.2'366.593  
DOI: 10.4312/linguistica.60.1.61-81



**Andrew Spencer\*\***  
University of Essex, UK

## **VOLITIONAL MOOD IN SOUTH SLAVIC WITH A FOCUS ON BULGARIAN: A PARADIGMATIC VIEW**

### **1 INTRODUCTION**

Most of the South Slavic languages have inflected imperative forms, as well as forms that seem to have the same directive force, but are analytic, rather than synthetic. In addition, there is a range of inflected and analytic forms which can be said to have related meanings, e.g. optative and hortative. We will follow the lead of Ammann and van der Auwera (2004) and use the term ‘volitional mood’ as a cover term for this set of related meanings and forms. Following their suggestion, the term ‘optative’ will be reserved for the expression of wishes, ‘imperative’ will be reserved for second person forms with directive force, and the term ‘hortative’ for appeals to act directed to first or third persons (with a further distinction between co-hortative for first person and ex-hortative for third persons possible). We return to the motivation for these distinctions below.

In this paper we investigate the relationship between inflected and analytic volitional mood forms, focusing in particular on analytic forms. We will argue that despite their syntactic nature, these forms exhibit some properties typical of paradigmatic organisation, e.g. complex interactions between different morphosyntactic features. In taking a paradigmatic approach, our paper builds on the review of syntactic volitional mood forms in Ammann and van der Auwera (2004), where they are seen as a features typical of Balkan languages. Here we suggest that analytic and inflected forms occupy the same information space and can be organised into a complex set of form and content paradigms (along the lines of Stump (2016)). We test these forms against the concept of periphrasis discussed in recent work in theoretical linguistics (Sadler/Spencer 2001; Ackerman/Stump 2004; Brown et al. 2012; Spencer/Popova 2015; Bonami 2015; Bonami et al. 2016; among others) and conclude that they exhibit some, but not all properties associated with canonical periphrases. In the next section we introduce inflected volitional mood and some analytical constructions in South Slavic, and argue that volitional mood paradigms tend to ‘fracture’, justifying the distinction between different sub-categories. Next we turn our attention to a particular analytic construction, the *da*-form construction, drawing primarily on data

---

\* g.popova@gold.ac.uk

\*\* spena@essex.ac.uk

from Bulgarian where it is well-developed, and argue that this polyfunctional form does not fill cells in an otherwise inflected paradigm, the way canonical periphrastic constructions do, but shares information space with inflected forms, exhibits non-compositionality and can be understood as a set of forms which are structured in paradigmatic oppositions with each other. We argue that this construction, whose volitional mood interpretation relies on it appearing in a main clause, should be constrained partially by syntax and partially by morphology. Finally, we sketch what an account along these lines might look like.

## 2 VOLITIONAL MOOD IN SOUTH SLAVIC

South Slavic languages have a set of synthetic inflectional forms with primarily imperative meaning, alongside, in most cases, a more or less developed paradigm of analytic forms with imperative, hortative and optative meanings. We provide a brief description of synthetic imperatives below, before turning to the analytic forms.

### 2.1 Synthetic Imperative/Hortative Forms

Like many other languages (for a very useful overview of imperatives see van der Wurff 2007), South Slavic synthetic imperatives have second person forms (singular and plural, also dual where dual verbal forms exist), which are used mostly in main imperative clauses. Inflectionally, these forms are fairly paired back: 2SG forms are equivalent to one of the stems of the verb, 2PL/DU forms add to the stem an ending indicating the respective person/number.<sup>1</sup> In addition, in some of the South Slavic languages, e.g. Slovene and Bosnian/Croatian/Serbian (BCS), there is also a synthetic 1PL (and 1DU where appropriate) form, which similarly adds the requisite person/number endings to the verbal stem. In (1) below we give some examples of synthetic inflected imperatives and first person hortatives across the South Slavic languages.

(1)	Slovene	BCS	Bulgarian	Macedonian
2SG	čitaj	čitaj	četi	čitaj
1DU	čitajva			
2DU	čitajta			
1PL	čitajmo	čitajmo		
2PL	čitajte	čitajte	četete	čitajte

Although part of the same paradigm, the first person forms have a somewhat different semantics, in that they express an appeal to do something to a group of addressees that includes the speaker and express an exhortation, rather than a command. As is

1 We use the following abbreviations 1/2/3 first/second/third person, ACC accusative, AOR aorist, CL clitic, DAT dative, DEF definite, DU dual, F feminine, FUT future, IMP imperative, IMPERF imperfect, LPTCP *l*-participle, M masculine, NOM nominative, OPT optative, PERF perfect, PFV perfective, PL plural, PROG progressive, PRS present, PST past, PTCP participle, REFL reflexive, SBJV subjunctive, SG singular.

clear from (1), not all the South Slavic languages have this form, for which Ammann and van der Auwera (2004) reserve the term (co-)hortative. This separation between the second person and first/third person forms is rooted partially in the inherent semantic distinctions between appeals to action made directly to an addressee, indirectly to a third person, or to self, possibly as a member of a group. This could, of course, be considered a polysemy, but another motivation for the distinction stems from intra- or inter-linguistic comparisons: different languages have forms that specialise for some, but not all of these meanings and we can find forms specialising for some of these meanings within the same language, as we shall see when we examine the analytic volitional mood forms. We might assume on these grounds that the forms in (1) really express two features: hortative and imperative.

The paradigms in (1) appear somewhat circumscribed, since both imperative and hortative cross-categorise with a limited number of the available feature-values in the language: they have limited person forms (there are no third person or 1sg forms in the paradigms shown above) and there are no tense distinctions, even though South Slavic languages have inflected tenses and, in some cases, for instance Bulgarian and Macedonian, a number of periphrastic tenses.

This is, however, not untypical of imperatives cross-linguistically (see van der Wurff 2007; Goussev 2013; among others). For this reason even in paradigmatic approaches to morphology the paradigms above are considered complete and the imperative paradigm is assumed a priori to be limited, for instance to second persons (see the treatment of the imperative in Bulgarian in Stump 2001, for instance). Inflected imperatives do not make tense distinctions and combine with limited person values. In this respect, however, they are fairly typical of this grammatical category (some tense marked imperative forms are noted in van der Wurff 2007, for instance, see also discussion of imperative paradigms in Goussev 2013).

There are some paradigm cells that could be considered gaps even in such circumscribed paradigms. Slavic verbs distinguish perfective and imperfective aspect and affirmative imperative forms are possible with both perfective and imperfective verbs. Negated imperative forms with perfective verbs, however, are either rarer (e.g. in BCS, see Szucsich 2010), or deemed impossible (for instance in Bulgarian). Descriptive grammars often motivate this gap in semantic terms. We return to this gap later.

## 2.2 Analytic Volitional Mood Constructions

Alongside the synthetic imperatives South Slavic languages also have a range of syntactic constructions with similar semantics, which in some cases complement and in others appear to be synonymous with the synthetic forms. What follows is not a comprehensive description of these constructions across all South Slavic languages. Instead, we will focus on one of the most widespread and productive ones – what we will call the *da*-verb forms – and will limit our discussion mostly to Bulgarian. Before we turn to our main data, however, we will show that some of the volitional mood constructions appear to support the distinction between different ‘sub-features’ under the label of ‘volitional mood’ put forward in Ammann and van der Auwera (2004). Heterogeneous

constructions some of which specialise for particular person-number combinations are not, it would appear, typologically surprising (see van der Auwera et al. 2004, for instance). Volitional mood constructions appear to be numerous, heterogeneous, and grammaticalised to a different degree (on the grammaticalisations of some volitional mood constructions in some Slavic languages see Hansen 2004 and Čakárova 2009, for example). There is a tendency, however, for some constructions to be restricted to certain person-number combinations, in other words, some constructions appear to reflect the distinctions between imperative (with second persons) and hortative (with first and/or third persons). For example, according to Herrity (2016), Slovene, which as we saw above has synthetic forms for the imperative and the co-hortative, has analytic constructions which we will refer to as ‘hortative’.<sup>2</sup>

According to Herrity (2016), these Slovene constructions are used to express a wish, desire, necessity, mild command, exhortation or regret on the part of the speaker. They are formed with the particle *naj* and either (i) 1st or 3rd persons of the present tense (ii) the future of the verb *bíti* ‘be’ or (iii) the third persons of the conditional (we illustrate (i) in 3 with examples from Herrity 2016, glosses added).

- (3) a. Vsè naj ostáne, kàkor je.  
 All let remain.3SG as be.3SG  
 ‘Let everything remain as it is.’
- b. Naj strokovnjáki réčejo, kàr hóčejo.  
 let experts say.3PL what want.3PL  
 ‘Let the experts say what they want.’

These forms are similar to the Bulgarian forms with the particle *neka* ‘let’, illustrated below (*neka* can optionally be followed by *da*):

- (4) a. Neka (da) idem na kino!  
 Let’s DA go.1PL to cinema  
 ‘Let’s go to the cinema.’
- b. Neka (da) se ugovorjat koga šte idem na kino.  
 Let DA REFL agree.3PL when FUT go.1PL to cinema  
 ‘Let them agree a time for us to go to the cinema.’

Although some scholars consider there to be a full paradigm of *neka* (*da*) forms, Ivanova and Gradinarova (2015: 56) note that in the contemporary language the second person forms are used very rarely.

2 Herrity (2016) refers to them as ‘optative mood’, a label we reserve here for forms that are used to express wishes, and notes that contemporary grammars of the language generally don’t recognise these forms as a distinct mood.



In BCS, the cognate constructions with *neka* seem to have specialised even further: Szucsich (2010) reports that in main clauses *neka* occurs primarily with third person forms. Other person forms seldom occur in main clauses. We give some of his examples (glosses adapted) for the *neka* construction in BCS in (5).

- (5) a. Neka pob(ij)edi naobolja ekipa.  
 OPT win.PRS.3.SG best.F.SG.NOM team.F.SG.NOM  
 ‘May the best team win.’
- b. Neka umrem <...>  
 OPT die.PRS.1.SG <...>
- otišla je drugom.  
 away.went.1LPTCP.F.SG be.3.SG other.M.SG.DAT  
 ‘Let me die ... she went to another one.’ (from a song text)

There are also constructions specialising for second persons, or imperatives. For example, in addition to the usual negation of imperatives with *ne*, there are specialised negated imperatives. In Bulgarian these are constructions with the special negative *nedej/nedejte*, historically (negated) imperative forms of the verb *dejati* ‘do’ and a *da*-form of the verb (or, less frequently, a remnant of the infinitive, which has practically disappeared from the language). The constructions with *da*-forms are illustrated below:

- (6) a. Nedej da xodiš na kino!  
 not.do.2SG DA go.2SG to cinema  
 ‘Don’t go (sg) to the cinema!’
- b. Nedejte da xodite na kino!  
 not.do.2PL da go.2PL to cinema  
 ‘Don’t go (pl) to the cinema!’

There are cognate negative imperative forms in other Slavic languages, see e.g. *nemoj/nemojte* constructions in BCS (Hansen 2004).

In contrast to these analytic constructions, which seem to have specialised or be specialising for particular persons-numbers, the constructions we will discuss in more detail, the *da*-form constructions, have forms for all person-number combinations. Though formally similar, in terms of meaning they can be organised around some of the different sub-features of the volitional mood. We turn to *da*-forms next.

### 3 DA-FORM VOLITIONAL MOOD CONSTRUCTIONS

*Da*-form constructions comprise the particle *da* and a verb inflected for tense and agreement features. The verb can be in some, though not all, of the tenses available in

the language. The particle/marker *da* is present throughout South Slavic. In Bulgarian, as in the other languages where it is found, it is polyfunctional and as a consequence its status is disputed and it has variably been considered a mood particle, or a complementiser; *da*-form constructions are also seen as a replacement of the infinitive, which some Slavic languages have (nearly) lost (a very useful overview can be found in Pitsch 2018). Pitsch (2018) summarises the arguments in favour of considering *da* a modal particle: unlike typical complementisers, *da* is strictly adjacent to the verb and can be separated from it only by clitics (like the negative particle *ne*, *da* behaves as a clitic which can head the clitic cluster); in addition, *da* can co-occur with undisputed complementisers like *če* ‘that’, but not with other modal particles like *bix* (conditional mood particle) and *šte* (the future tense particle). In the remainder of this paper we will adopt this position and will treat *da* as a verbal/modal particle. The modality associated with *da* in most of its uses comes from its tendency to appear in irrealis contexts.

Certain *da*-forms are traditionally included in grammatical descriptions of Bulgarian (e.g. Nicolova 2008 and references therein) as alternative or complementary to the inflected imperative forms see (7), in other words, these forms have volitional mood meanings. We illustrate the basic constructions with the perfective verb *kupja* ‘buy’, but *da*- constructions can be used with both perfective and imperfective verbs.

(7)

	<i>da</i> +present tense	<i>da</i> +present perfect	<i>da</i> +past perfect	<i>da</i> +imperfect
1SG	da kupja	da sām kupil/a/o	da bjax kupil/a/o	da kupex
2SG	da kupiš	da si kupil/a/o	da beše kupil/a/o	da kupeše
3SG	da kupi	da e kupil/a/o	da beše kupil/a/o	da kupeše
1PL	da kupim	da sme kupili	da bjaxme kupili	da kupexme
2PL	da kupite	da ste kupili	da bjaxte kupili	da kupexte
3PL	da kupjat	da sa kupili	da bjaxa kupili	da kupexa

Volitional mood meanings for the *da*-forms illustrated above are available when they are used in main clauses, compare (8) with (9). The phenomenon is described and discussed as a possible Balkanism in Ammann and van der Auwera (2004). Some examples of the uses of such main clause *da*-constructions are given in (8).

- (8) a. Da kupiš xľjab!  
 DA buy.2SG.PRS bread  
 ‘Buy bread!’
- b. Da imaš mnogo kšmet!  
 DA have.2SG.PRS much luck  
 ‘May you have a lot of luck!’

- c. Do utre da sme kupili xljab!  
 by tomorrow DA be.3PL buy.LPTCP.PL bread  
 ‘We must have bought bread by tomorrow/Let’s buy bread by tomorrow.’
- d. Ex, da bjax kupila xljab!  
 ah DA be.IMPERF.1SG buy.LPTCP.3SG.F bread  
 ‘I wish I had bought bread.’
- e. Ex, da imax sega mnogo pari!  
 ah DA have.IMPERF.1SG now much money  
 ‘If only I had a lot of money!’

The same *da*-forms are found in subordinate clauses, indeed the typical uses of *da*-forms, usually considered analogues of subjunctives and/or infinitives, are in subordinate clauses. The *da*-forms with present tense verb are the most extensively discussed in the literature and seem to have the widest range of uses. We illustrate some of this range in (9), where subordinate *da*-clauses are used as complements of control verbs (9a, b, c), purpose clauses (9d), as well as subject clauses (9e) and nominal complements (9f).

- (9) a. Naredix da kupiš xljab.  
 order.1SG.PST DA buy.PRS.2SG bread.  
 ‘I ordered (you) to buy bread.’
- b. Naredix Ivan da kupi xljab.  
 order.1SG.PST Ivan DA buy.PRS.3SG bread.  
 ‘I ordered Ivan to buy bread.’
- c. Togava šte započneš da peeš.  
 then FUT start.2SG DA sing.PRS.2SG  
 ‘Then (you) will start singing.’
- d. Zaminax da uča v čužbina.  
 depart.AOR.1SG DA study.PRS.1SG in abroad  
 ‘I left to study abroad.’
- e. Da kupiš xljab e neobxodimo.  
 DA buy.PRS.3SG bread is.3SG necessary  
 ‘(For you) to buy bread is necessary. (It is necessary for you to buy bread.)’
- f. Tja ima namerenie da uči v čužbina.  
 she has intention DA study.PRS.3SG in abroad  
 ‘She has the intention to study abroad.’

*Da*-forms with other tenses are also possible in subordinate clauses. *Da*-forms with the perfect tense are used in irrealis future-perfect contexts (10), whereas *da*-forms with the imperfect and past perfect are found in the protasis of conditional sentences with past semantics (11), and in other subordinate clauses with counterfactual meaning (12).

(10) Nadjavam se do do večera  
 hope.PRS.1SG REFL by tonight  
 da ste kupili xljab.  
 DA be.PRS.3SG buy.LPTCP.PL bread  
 ‘I hope that by tonight you will have bought bread.’

(11) a. Da bjaxte stojali na opaški pred  
 DA be.IMPERF stand.LPTCP on queues in-front-of  
 izbiratelnite urni, sega štxajte da bādete evropejci.  
 ballot.DEF boxes now want.IMPERF DA be.3PL Europeans  
 ‘Had you queued in front of the ballot boxes, you would have been Europeans now.’

b. Dori da imaxme obxvat, ne bixme  
 even DA have.IMPERF.1PL coverage not be.COND.1PL  
 mogli da pomognem koj znae kolko.  
 be.able DA help.PRS.1PL who knows how.much  
 ‘Even if we had coverage, we wouldn’t be able to help much.’

(12) Knigata, kojato iskate roditelite vi  
 book.DEF which want.PRS.2PL parents.DEF your.CL  
 da bjaxa pročeli, veče e na pazara.  
 DA be.IMPERF.2PL read.LPTCP.PL already is.PRS.3SG on market.DEF  
 ‘The book you wish your parents had read is already on the market.’

*Da*-forms in main clauses are not exclusive to Bulgarian, they occur in South Slavic more generally. We list some examples below, though given the rather complex nature of the phenomenon, we don’t undertake a detailed comparison.

For BCS, Alexander (2006: 249–250) points out that a *da*-clause that stands alone can communicate a wish, desire or request. We provide some of her examples in (13), with added glosses. Independent *da*-clauses seem to be possible with both present and past tense, with past tenses they are said to express a veiled threat.

(13) a. Da vas upoznam.  
 DA you.PL introduce.PRS.1SG  
 ‘Let me introduce you.’

- b. Da se nagodimo.  
 DA REFL agree.PRS.1PL  
 ‘Let’s make a deal.’
- c. Samo da daska ne pukne!  
 only DA board not break.PRS.3SG  
 ‘If only the board doesn’t break! (I hope that the board won’t break!)’
- d. Da to više nikad nisi učinio!  
 DA that more never not.BE.2SG do.PTCP  
 ‘Don’t you ever do that again!’

Uses of *da*-clauses with volitional meanings in BCS are described also in Szucsich (2010). A comparison between *da* and *neka* highlights the propensity of *da* to appear in irrealis contexts. Uses of *da* with the subjunctive are also recorded. We give some of Szucsich’s (2010: 400ff.) examples (with adapted glosses) in (14).

- (14) a. Samo da bi došao!  
 only that SBJV.3SG come.LPTCP.M.SG  
 ‘If only he came!’
- b. Da si samo brže vozio!  
 that be.2SG only faster drive.LPTCP.M.SG  
 ‘If only you had gone faster!’
- c. Da ste odmah išli u prodav(ao)nicu!  
 that be.2.PL immediately go.LPTCP.M.PL in store.F.SG.ACC  
 ‘Immediately, go to the store!’

In Macedonian, Mišeska Tomić (2012: 374–375) notes the use of a *da*-form with a perfective imperfect verb to express a missed opportunity, her example is reproduced in (15) (glosses adapted).

- (15) a. Da dojde!  
 DA come.3PL.PFV.IMPERF  
 ‘They (were invited and) should have come! (They shouldn’t blame anybody for not coming.)’
- b. Da go zemeše!  
 DA 3SG.M.ACC.CL take.2/3SG.PFV.IMPERF  
 ‘You/(s)he should have taken it! (Now you/(s)he cannot blame anyone for not taking it.)’

The phenomenon of volitional mood main clause constructions appears to be an example of what Evans (2007) calls ‘insubordination’, the use of typical subordinate clauses as main clauses, sometimes with conventionalised semantics. In the case of the South Slavic constructions this conventionalisation of insubordination is expressed in the association of *da*-form main clause uses with volitional mood semantics. As a result, these *da*-form uses are semantically non-compositional, in other words, the imperative meaning associated with *da kupiš (x)ljab* ‘to buy bread’ comes precisely from the use of the construction in a main clause, rather than from the meaning of any of its components, which can appear in different syntactic contexts with different interpretations.

### 3.1 The Content Paradigm of Analytic Volitional Mood Constructions

In a number of ways volitional mood *da*-forms are similar to other analytic forms that have been considered periphrastic, i.e. syntactic forms that bear similarities to inflected word-forms, discussed recently in a score of studies (e.g. Sadler and Spencer 2001; Ackerman/Stump 2004; Brown et al. 2012; Bonami/Webelhuth 2013; Bonami 2015; Spencer/Popova 2015; Bonami et al. 2016; among others). The most clear-cut definitional property for periphrasis is paradigm intersection: the case where a syntactic construction occupies a cell in an otherwise inflected paradigm. We will discuss a putative case of paradigm intersection between the inflected imperative and the analytic *da*-volitional forms below, but will conclude that although there are complex interactions between the inflected and the syntactic paradigm, it is difficult to claim a straightforward case of paradigm intersection. However, there is another property – (morphosyntactic) non-compositionality – that has been put forward as definitional for periphrasis in Ackerman and Stump (2004), for instance, see also discussion in Brown et al. (2012). As we saw before, *da*-form volitional constructions are non-compositional. We will therefore treat them as periphrastic, i.e. as forms that are, on some sense, equivalent to inflected forms of lexemes, which can be seen as part of the paradigm of these lexemes and which can, at least in part, be regulated by the morphology. We will sketch below what form this interaction could take.

The *da*-clauses are polyfunctional in much the same way that inflected word forms can be polyfunctional, a situation described in morphology as syncretism. For instance, the *-ing* form of an English verb is ambiguous between four or five separate uses, as the progressive aspect converb (*Mary is singing*), an adjectival, participial form (*the girl singing the song last night*), a gerund or adverbial usage (*Singing loudly, Mary walked on stage*) and perhaps two distinct types of nominalization (*We didn't expect you singing so loudly; Your singing is loud*). In the inferential-realizational morphological model advocated in Stump (2016), for instance, such syncretisms are handled by setting up two types of paradigm for an inflecting lexeme, content paradigms and form paradigms. The content paradigm is defined over the set of morphosyntactic properties accessible to syntax and semantics. In the case of English verbs this would include the features defining the four or five uses just illustrated (say, [ASPECT: prog], [PTCP: prs] and so on). The feature value pairs would thus define four or five distinct cells in

the content paradigm of the lexeme *SING*. The form paradigm, on the other hand would be much simpler and would include just one cell corresponding to the four of five cells of the content paradigm: [Vform: ing]. In Stump's model the content and form paradigms are related through Paradigm Linkage, mediated by a function which in the present case would map each of the content paradigm cells realizing [ASPECT: prog], [PTCP: prs] etc. to the single form paradigm cell [Vform: ing].

We propose that, similar to the synthetic imperatives, Bulgarian *da*-forms in (7) can be interpreted as a form paradigm. It maps from content paradigms in complex ways. Using available descriptions of these forms, below we sketch a possible content-paradigmatic organisation of these constructions, based on the distinctions we introduced earlier.

Some *da*-forms can be used as synonyms of the synthetic imperative – closest in meaning are the second person forms with the present tense. We illustrated with (8a) for the second singular, repeated here.

- (8) a. [repeated]  
 Da kupaš xljab!  
 DA buy.2SG.PRS.PFV bread  
 'Buy bread!'

In paradigmatic terms, forms like (8a) and the analogous 2PL constructions create an over-abundance in the system. However, the *da*-construction forms, unlike the synthetic imperatives, cross-categorise with all person-numbers and also interact with tense, as the labels we have been using and the examples in (7) suggest. These are, however, what might be called non-cumulative paradigms, in other words, the present tense *da*-forms are not simply forms in which the semantics of the present tense is added to some other semantics marked elsewhere in the construction. On the contrary, the tense semantics is neutralised (similar to the synthetic imperatives, the present tense forms have futurate interpretation) Since these forms are very close in meaning to the synthetic imperative, we will call them 'unmarked'. It should be noted, however, that the present tense 2sg and 2pl *da*-forms are less formal and have stronger pragmatic force than their synthetic counterparts (see Nicolova 2008 for details). The present tense *da*-forms can also be used as hortative 1PL<sup>3</sup> (see 16) and third person forms (see 17).

- (16) Da kupim xljab!  
 DA buy.1PL bread  
 'Let's buy bread.'

- (17) a. Deteto da večerja!  
 child.DEF DA eat.dinner.3SG  
 'The child should have dinner.'

---

3 1SG forms are rare, as one would expect of appeals to act addressed to self.

- b. Decata da večerjat!  
 children.DEF DA eat.dinner.3PL  
 ‘The children should have dinner.’

The perfect tense *da*-forms can be used as commands/exhortations which reference an action that should be completed by some future point in time (compare 18a with 18b).

- (18) a. Do dovečera da si kupil xljab!  
 by tonight DA be.2SG buy.LPTCP.M bread  
 ‘(You) should have bought bread by tonight.’

- b. Do dovečera šte si kupil xljab.  
 by tonight FUT be.2SG buy.LPTCP.M bread  
 ‘(You) will have bought bread by tonight.’

The use of the 1PL and third person perfect tense *da*-forms is similar. Thus, the imperative-hortative paradigm of *da*-forms has two cells: an ‘unmarked’ one (the present tense *da*-forms) and a ‘future perfect’ one (perfect tense *da*-forms).

The present tense *da*-forms can be used also for wishes and curses, i.e. to express the optative mood in the terminology of Ammann and van der Auwera (2004). One set of optative mood forms is syncretic to the unmarked imperative/hortative, i.e. these are the present tense *da*-forms. In (19) we illustrate with the wish ‘may (I/you/she/he/it/ etc.) be healthy’:

- |      |                    |                  |
|------|--------------------|------------------|
| (19) | SG                 | PL               |
| 1    | da bāda zdrav/a/o  | da bādem zdravi  |
| 2    | da bādeš zdrav/a/o | da bādete zdravi |
| 3    | da bāde zdrav/a/o  | da bādat zdravi  |

This use is available to synthetic volitional mood forms as well, but only with the second person (20).

- |      |                |                  |
|------|----------------|------------------|
| (20) | SG             | PL               |
| 2    | bādi zdrav/a/o | bādete zdrav/a/o |

The optative can be expressed also by *da*-forms with imperfect or past perfect in main clauses. These forms are interpreted as counterfactual wishes, see (21). In other words, this part of the paradigm also has two cells: unmarked optative and counterfactual optatives.



- (21) a. Ex, da bjaxme kato Estonia!  
 Ah, DA be.IMPERF.1PL like Estonia  
 ‘If only we were like Estonia!’
- b. Ex, da imax pari!  
 Ah, DA have.IMPERF.1SG money  
 ‘If only I had money!’ or ‘I wish I had money!’
- c. Pone lipite da bjaxa požalili!  
 at.least linden.trees DA be.IMPERF take.pity.LPTCP.PL  
 ‘(I wish) they had at least spared the linden trees!’
- d. Da bjaxa kupili xljab!  
 DA be.IMPERF.3PL buy.LPTCP.PL bread  
 ‘I wish they had bought bread!’
- e. Da bjaxa imali kăsmet da idat na more!  
 DA be.IMPERF.3PL have.LPTCP.PL luck DA go to sea  
 ‘I wish they had the luck to go to the seaside.’

To sum up, what we find in Bulgarian is a set of *da*-forms with verbs in the present, perfect, past perfect and imperfect tenses, which map onto functions in a complex way. They partially overlap with synthetic imperative forms, but their own paradigm is fuller. In some cases the *da*-forms have meanings very similar to the meaning of the inflected imperative, but as there are stylistic differences between them, we will consider them almost synonymous forms that nonetheless belong to two separate, if semantically related paradigms.

There is one set of forms that could be seen to fill in a gap in the inflected paradigm, and these are the negated perfective *da*-forms. Inflected imperative forms, just like *da*-forms, are negated with the default negator *ne*. In some South Slavic languages (Slovene, for instance) negated inflected imperatives are possible with both imperfective and with perfective verbs. Negated imperatives with perfective verbs have a slightly different semantics, sometimes denoted with the term preventative – for warnings, or when the speaker wishes to draw attention to the unwanted consequences of a particular action. We illustrate with a Slovene example from Herryty (2016) in (22), glosses added.

- (22) Ne odpri vrát za nobêno céno  
 not open.IMP.2SG.PFV door for no price  
 ‘Do not open the door on any account.’

In Bulgarian synthetic imperatives are disallowed with perfective verbs when negated. With volitional *da*-forms, on the other hand, there is no such restriction. When

perfective *da*-forms with imperative function are negated, they also function as a preventative imperative, i.e. they are appropriate in those cases where a speaker considers an event to be undesirable and urges the listener to take action to prevent that event from happening (Academy Grammar, vol. 3, 68, see examples in 23).

- (23) a. Da ne padneš!  
da not fall.2SG  
'(be careful) Don't fall!'
- b. Da ne nastinete!  
da not catch.cold.2PL  
'(be careful) Don't catch a cold!'

Although in Bulgarian the *da*-forms have taken on a function which is performed by the inflected imperative in other languages, it is difficult to consider this a gap in the paradigm because of the 'fractured' nature of the volitional mood paradigms, that is, the propensity of different person-number combinations in this paradigm to assume a different meaning, i.e. the propensity of the form paradigms to 'fracture' into a number of separate paradigms. Although we will consider the inflectional and the syntactic paradigm independent of each other, we will suggest below that we can still make use of the fact that they occupy a similar information space to block inflected negated imperatives with perfective verbs in Bulgarian.<sup>4</sup>

#### 4 ACCOUNTING FOR ANALYTIC VOLITIONAL MOOD CONSTRUCTIONS: A SKETCH OF A PROPOSAL

In the previous sections we argued that *da*-form constructions, when used in main clauses, play a role in the language which is very similar to that of inflected imperative forms. Volitional mood *da*-form constructions can also be seen as a paradigm, which depends on the cross-categorisation of mood and tense features. At the same time *da*-form constructions are similar syntactically to *da*-forms used in subordinate clauses. Here we propose that this complex behaviour calls for an account in which the properties of these constructions are constrained partially by the syntax and partially by the morphology. Our account is thus similar to the account we have proposed for periphrastic constructions in other work. What *da*-forms show even more clearly than more canonical periphrases is that grammatical analytic constructions share properties with other syntactic constructions in the language, while at the same time expressing information which elsewhere can be expressed by indisputably inflected forms (a point made as early as the seminal Ackerman/Webelhuth 1998).

---

4 For a fuller overview of this in Slavic and for an account of the Bulgarian data which attributes the unavailability of inflected imperative with perfective verbs in Bulgarian to the morphosemantics of tense and aspect categories in the language, see Kuehnast (2008). Our account is closer to the traditional view that the loss of these forms became more likely due to the existence of synonymous syntactic constructions.

The analysis sketched below is based on the assumptions of a paradigm-based approach to morphology along the lines of Paradigm Function Morphology (PFM) (Stump 2001, and later updates of the framework as in Stump 2016, see also Bonami/Stump 2016), on the one hand, and lexicalist syntactic frameworks like Head-driven Phrase Structure Grammar (HPSG) (Pollard/Sag 1994, especially HPSG approaches that integrate constructions Ginzburg/Sag 2000). PFM is an inferential-realizational framework in which the phonological form of an inflected word-form is inferred from the information associated with the cell in the paradigm of that form. HPSG approaches like those in Sag (1997) and Ginzburg and Sag (2000), on the other hand, assume that the grammar contains descriptions of phrasal signs arranged in a default inheritance hierarchy, much like descriptions of words can be. This allows phrasal signs to share properties or be more or less specific than each other. Some constraints over phrases are default constraints and can be overridden when stated over more specific constructions.

We will assume that *da*-forms are constructions in this sense, i.e. they are phrasal signs, partially constrained by their descriptions in the grammar. In some constructional approaches within HPSG certain constraints are specified to be default constraints, i.e. they can be overridden in specific constructions. This is the case with the Head-Feature Principle, which says that the HEAD features of a phrasal sign with a head-daughter and a complement daughter would be inherited from the head-daughter. In what follows we assume that HEAD features are not necessarily passed on to the phrase level by the head-daughter, where a phrasal description or some other constraints says otherwise.

As we said earlier, we assume that *da* is a modal particle and we account for it as a raising verb, much like the account of English infinitival *to* in Sag et al. (2003), for instance. We show the properties of a subordinate *da*-construction in Figure 1. For simplicity, we focus on constructions with present tense verbs.

This construction has a head daughter (HD-DTR) – *da* itself – and one other daughter on its daughters list (DTRS) – the lexical verb. The description of this construction specifies that the lexical verb should be in the present tense, and that its agreement and aspectual values are shared with the construction as a whole (this sharing is indicated by numbered tags like [4]). The construction itself is specified to have VFORM *daform*, to allow it to be subcategorized for by verbs that require *da*-forms as their complements.

In this construction, the present tense is not passed on to the level of the construction. This reflects the observation that in at least some *da*-form constructions the present tense is pleonastic (see discussion in Pitsch 2018). The subject [s] on the list of arguments of the lexical verb daughter is also the subject of the HD-DTR. The construction as a whole inherits its lexemic identity (LID) from the lexical verb. An important specification on the level of the construction is the negative value for the feature I(independent) C(lause) (a similar feature is used in Ginzburg/Sag 2000). This ensures that this construction cannot be used as a main clause.

Volitional mood *da*-form constructions are very similar to subordinate constructions. The information they express, however, is partially specified by the morphology. In some formalisations of PFM-style approaches word-forms are constrained by implicational statements whose antecedent mentions relevant morphosyntactic properties of

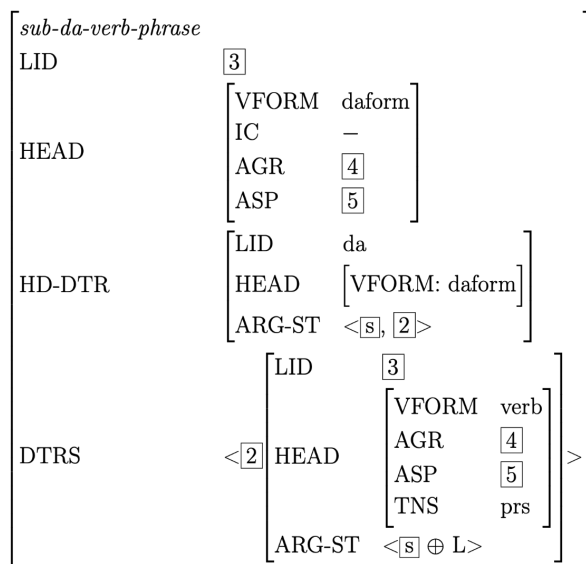


Figure 1: The *da*-construction

signs and whose consequent states properties of words that can express these morpho-syntactic properties (see Crysmann/Bonami 2016). Such implicational statements are in Paninian competition – those with a more specific set of morphosyntactic properties in the antecedent ‘win out’, i.e. pre-empt the application of implicational statements with more general antecedents. The paradigm cells of the analytic volitional mood constructions include the specification VFORM *da*-form alongside features associated with the mood they express, for example [MOOD: *imp*]. This feature would appear in the relevant antecedents and ensure that the syntactic paradigms don’t directly compete with the inflected imperative paradigm. The imperative main clause syntactic constructions, for instance, can be subject to the constraints in (25).

$$(25) \quad \left[ \text{HEAD} \left[ \begin{array}{l} \text{MOOD} \quad \textit{imp} \\ \text{VFORM} \quad \textit{daform} \end{array} \right] \right] \rightarrow \left[ \begin{array}{l} \textit{in-sub-da-verb-phrase} \\ \text{HEAD} \left[ \begin{array}{l} \text{MOOD} \quad \textit{imp} \\ \text{VFORM} \quad \textit{daform} \\ \text{IC} \quad + \end{array} \right] \end{array} \right]$$

This constraint specifies that the cell in the paradigm that includes the features [MOOD *imp*] and [VFORM *daform*] is realised by a sub-type of *da*-constructions which has as part of its head-features the specification [MOOD *imp*], [VFORM *daform*] and a positive value for the feature IC, in other words, this construction has to appear in a main clause.

A fuller description of the construction is given in Figure 2.

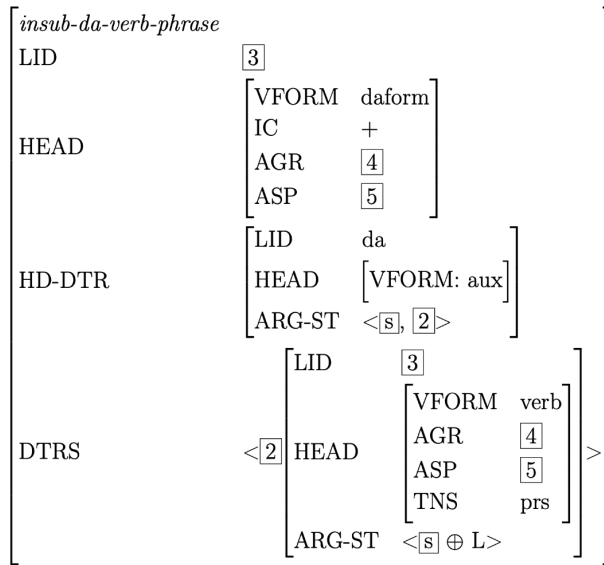


Figure 2: The main-clause imperative *da*-construction

And finally, we can capitalise on the fact that there are no negated inflected imperatives with perfective verbs, we could assume the constraint in (26), which gives as the realisation of the respective paradigm cell the *da*-form construction, rather than an inflected verb.

$$(26) \quad \left[ \begin{array}{l} \text{HEAD} \left[ \begin{array}{l} \text{MOOD} \quad \text{imp} \\ \text{ASP} \quad \text{pfv} \\ \text{NEG} \quad + \end{array} \right] \end{array} \right] \rightarrow \left[ \begin{array}{l} \textit{in-sub-da-verb-phrase} \\ \text{HEAD} \left[ \begin{array}{l} \text{MOOD} \quad \text{imp} \\ \text{ASP} \quad \text{pfv} \\ \text{NEG} \quad + \\ \text{VFORM} \quad \text{daform} \\ \text{IC} \quad + \end{array} \right] \end{array} \right]$$

## 5 CONCLUSION

As the preceding discussion suggests, the volitional mood paradigm is ‘fractured’: we find somewhat different, but related, meanings for specific cells in the paradigm. This makes it difficult to claim that a volitional mood paradigm is ‘deficient’, since such deficiency is not exceptional: the imperative hierarchy formulated in van der Auwera et al. (2004), for instance, predicts that cross-linguistically different person-number combinations may be associated with different forms, or indeed, with no forms at all. Thus, in South Slavic we find a somewhat sparse but typical inflected paradigm. However, in some languages, for example Bulgarian, the volitional mood is also associated

with a rather rich range of analytic forms. We focused on one set of such forms, the constructions with the modal particle *da*. We claimed that these form a syntactic paradigm, part of the volitional mood category. Given the non-compositionality of this construction, the fact that it shares properties with other constructions in the language and that it shares information space with the inflected imperatives we suggested that it should be constrained jointly by the morphology and by the syntax and outlined a tentative analysis.

## References

- Academy Grammar, vol. 3 (1983) *Gramatika na sǎvremenniiia bǎlgarski knižoven ezik. Sintaksis*. Sofia: Izdatelsto na Bǎlgarskata akademija na naukite.
- ACKERMAN, Farrell/Gregory STUMP (2004) “Paradigms and periphrastic expressions.” In: L. Sadler/A. Spencer (eds), *Projecting morphology*. Stanford Studies in Morphology and the Lexicon. Stanford, California: CSLI Publications, 111–157.
- ACKERMAN, Farrell/Gert WEBELHUTH (1998) *A theory of predicates*. Stanford, CA: CSLI Publications.
- ALEXANDER, Ronelle (2006) *Bosnian, Croatian, Serbian: A grammar with sociolinguistic commentary*. Madison, WI: The University of Wisconsin Press.
- AMMANN, Andreas/Johan VAN DER AUWERA (2004) Complementizer-headed main clauses for volitional moods in the languages of South- Eastern Europe: A Balkanism? In: O. Mišeska Tomić (ed.), *Balkan syntax and semantics, vol. 67 Linguistik Aktuell/Linguistics Today*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 293–314. <https://doi.org/10.1075/la.67>
- BONAMI, Olivier (2015) “Periphrasis as collocation.” *Morphology* 25/1, 63–110. <https://doi.org/10.1007/s11525-015-9254-3>
- BONAMI, Olivier/Robert D. BORSLEY/Maggie TALLERMAN (2016) “On pseudo-non-finite clauses in Welsh.” In: D. Arnold/M. Butt/B. Crysmann/T. Holloway King/S. Müller (eds), *Proceedings of the joint 2016 conference on Head-driven Phrase Structure Grammar and Lexical Functional Grammar. Polish Academy of Sciences, Warsaw, Poland*. Stanford, CA: CSLI Publications, 104–124.
- BONAMI, Olivier/Gregory STUMP (2016) “Paradigm function morphology.” In: A. Hippisley/G. Stump (eds), *Cambridge Handbook of Morphology*. Cambridge University Press, 449–481.
- BONAMI, Olivier/Gert WEBELHUTH (2013) “The phrase-structural diversity of periphrasis: a lexicalist account.” In: M. Chumakina/G. Corbett (eds), *Periphrasis: The role of syntax and morphology in paradigms*. Oxford: Oxford University Press, 141–167.
- BROWN, Dunstan/Marina CHUMAKINA/Greville CORBETT/Gergana POPOVA/Andrew SPENCER (2012) “Defining ‘periphrasis’: key notions.” *Morphology* 22, 233–275. <https://doi.org/10.1007/s11525-012-9201-5>
- ČAKĀROVA, Krasimira. 2009. *Imperativāt v sǎvremennija bǎlgarski ezik*. Plovdiv: Izdatelstvo “Pigmalion”.

- CRYSMANN, Berthold/Olivier BONAMI (2016) “Variable morphotactics in Information-based Morphology.” *Journal of Linguistics* 52/2, 311–374. <https://doi.org/10.1017/S0022226715000018>
- EVANS, Nicholas (2007) “Insubordination and its uses.” In: I. Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, 366–431. Oxford: Oxford University Press.
- GINZBURG, Jonathan/Ivan A. SAG (2000) *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford, CA: CSLI Publications.
- GOUSSEV, V. Ju. (2013) *Tipologija imperativa*. Moscow: Jazyki slavjanskoj kul'tury.
- HANSEN, Björn (2004) “The grammaticalization of the analytical imperatives in Russian, Polish and Serbian/Croatian.” *Die Welt der Slaven* XLIX/2, 257–274.
- HERRITY, Peter (2016) *Slovene: A comprehensive grammar*. London and New York: Routledge.
- IVANOVA, E. Ju./A. A. GRADINAROVA (2015) *Sintaksičeskaja sistema bolgarskogo jazyka na fone russkogo*. Moskva: Jazyki slavjanskoj kul'tury.
- KUEHNAST, Milena (2008) “Aspectual coercion in Bulgarian negative imperatives.” In: W. Abraham/E. Leiss (eds) *Modality-Aspect interfaces: Implications and typological solutions*. Amsterdam: John Benjamins Publishing Company, 175–196.
- MIŠESKA TOMIĆ, Olga (2012) *A grammar of Macedonian*. Bloomington, In.: Slavica Publishers.
- NICOLOVA, Ruselina (2008) *Bălgarska gramatika: Morfologija*. Sofija: Universitetsko izdatelstvo “Sv. Kliment Oksridski”.
- PITSCH, Hagen (2018) “Bulgarian moods.” *Journal of Slavic Linguistics* 26/1, 55–100. <https://doi.org/10.1353/jsl.2018.0003>
- POLLARD, Carl/Ivan A. SAG (1994) *Head-driven phrase structure grammar*. Chicago: University of Chicago Press.
- SADLER, Louisa/Andrew SPENCER (2001) “Syntax as an exponent of morphological features.” In: G. Booij/J. van Marle (eds) *Yearbook of morphology 2000*. Dordrecht: Kluwer, 71–96.
- SAG, Ivan A. (1997) “English relative clause constructions.” *Journal of Linguistics* 33, 431–484.
- SAG, Ivan A./Thomas WASOW/Emily M. BENDER (2003) *Syntactic theory: A formal introduction*, 2<sup>nd</sup> edn. Stanford, CA: CSLI Publications.
- SPENCER, Andrew/Gergana POPOVA (2015) “Periphrasis and inflection.” In: *The Oxford Handbook of Inflection*. Oxford: Oxford University Press, 197–230.
- STUMP, Gregory (2001) *Paradigm Function Morphology. A theory of inflectional form*. Cambridge: Cambridge University Press.
- STUMP, Gregory (2016) *Inflectional Paradigms: Content and form at the syntax-morphology interface*. Cambridge: Cambridge University Press.
- SZUCSICH, Luka (2010) “Mood in Bosnian, Croatian and Serbian.” In: B. Rothstein/R. Thieroff (eds) *Mood in the Languages of Europe, vol. 120 Studies in Language Companion Series (SLCS)*. Amsterdam/Philadelphia: John Benjamins Publishing, 394–408.

- VAN DER AUWERA, Johan/Nina DOBRUSHINA/Valentin GOUSSEV (2004) "A semantic map for imperative-hortatives." In: D. Willems/B. Defrancq/T. Colleman/D. Noël (eds), *Contrastive Analysis in Language: Identifying linguistic units of comparison*. London: Palgrave Macmillan, 44–66.
- VAN DER WURFF, Wim (2007) "Imperative clauses in generative grammar. An introduction." In: W. van der Wurff (ed.), *Imperative Clauses in Generative Grammar: Studies in honour of Frits Beukema*. Amsterdam/Philadelphia: John Benjamins Publishing, 1–94.

#### Abstract

### VOLITIONAL MOOD IN SOUTH SLAVIC WITH A FOCUS ON BULGARIAN: A PARADIGMATIC VIEW

In this paper, we survey some of the inflected and periphrastic volitional mood paradigms in South Slavic with a focus on Bulgarian data. Our review confirms typological observations in the literature that volitional mood paradigms tend to 'fracture', in that the cross-categorisation with different person/number features leads to systematic associations with different meanings and, typologically, frequent associations with different forms. This makes it difficult to argue that non-existent inflected forms are gaps in the inflected paradigm. Because of this periphrastic volitional mood forms that exist alongside inflected forms have to be seen as independent (syntactic) paradigms, rather than forms that fill missing cells in inflected paradigms, i.e. the product of feature intersection. Like more canonical periphrases, however, syntactic volitional mood forms are non-compositional and exhibit an organisation akin to the content-form paradigm organisation of inflected paradigms developed in certain inferential-realizational approaches to morphology. Following some recent formalisations, we suggest a tentative analysis of the most productive periphrastic volitional mood forms in Bulgarian based on the assumption that their properties are constrained partially by the morphology and partially in the syntax.

**Keywords:** Bulgarian, Slavic, volitional mood, periphrasis

#### Povzetek

### VELELNI NAKLON V JUŽNOSLOVANSKIH JEZIKIH S POUĐARKOM NA BOLGARŠČINI: PARADIGMATSKI POGLED

Prispevek obravnava pregibne in sestavljene paradigme velelnega naklona v južnoslovanskih jezikih s poudarkom na podatkih iz bolgarščine. Pregled literature potrđi tipološka opažanja o 'lomljenju' paradigmem velelnega naklona, saj se navzkrižna kategorizacija oznak osebe/števila sistematično povezuje z različnimi pomeni in, s tipološkega vidika, pogosto tudi z različnimi oblikami. Zato težko trdimo, da neobstoječe



pregibne oblike predstavljajo vrzeli v pregibni paradigmi. Sestavljene oblike velelnega naklona, ki obstajajo ob pregibnih, torej razumemo kot samostojne (skladenjske) paradigme, in ne kot oblike, ki zapolnijo prazna mesta v pregibnih paradigmah, tj. rezultat križanja oznak. Kot bolj kanonične sestavljene oblike pa so skladenjske oblike velelnega naklona nekompozicijske in izkazujejo organizacijo, ki je podobna tisti v paradigmi vsebina-oblika pri pregibnih paradigmah, kot jih poznajo v nekaterih inferenčno-realizacijskih pristopih k morfologiji. V prispevku sledimo novejšim formalnim pristopom in predlagamo razčlenbo najbolj produktivnih sestavljenih oblik velelnega naklona v bolgarščini, ki privzema, da so lastnosti teh oblik omejene tako oblikoslovno kot skladenjsko.

**Ključne besede:** bolgarščina, slovanski jeziki, velelni naklon, sestavljene oblike





## OV IS IN THE AIR: THE EXTREME MULTIFUNCTIONALITY OF THE SLOVENIAN AFFIX OV<sup>1</sup>

### 1 INTRODUCTION

One of the classical features of Distributed Morphology (Halle/Marantz 1993, 1994) is the assumption that derivational affixes correspond to categorial heads (n, v, a). However, several recent proposals within Distributed Morphology (Lowenstamm 2014; Nevins 2015; Creemers et al. 2017) argue that (either some or all) derivational affixes should be analyzed as roots. As core evidence, these authors discuss affixes which are promiscuous both in terms of what categories they select and what categories they can realize (or, be selected by, if they are viewed as roots). We observe that such promiscuous affixes discussed in the literature have either highly abstract or highly unpredictable meaning, potentially pointing in the direction of no stored meaning at all. Working in a related approach, Simonović and Arsenijević (2020) argue that certain affixes in Serbo-Croatian function either as inflectional or as derivational, deriving either passive participles or adjectives. Importantly, in this case as well, once the categorial content is removed, the affix/root appears rather light in terms of its semantic content.

Testing the limits of the unification of affixes which appear both in inflection and derivation and with different categorial embeddings, we focus on the Slovenian affix *ov*, which can be found in nouns, adjectives and verbs. Based on this distribution, one could at first sight take these occurrences of *ov* as several different affixes which coincidentally have the same phonological form. However, considering the semantic contribution of *ov* as well as its prosodic effects, we will argue for a single, extremely multifunctional *ov*. As a result, we will provide a comprehensive analysis of the morpheme *ov* in Slovenian and give further support for the unification approach of affixes proposed in the literature. Furthermore, in dealing with the specific contexts where *ov* appears, we further elaborate the model initially proposed

\* rkiema@gmail.com

\*\* petra.mismas@ung.si

1 We acknowledge the financial support by the Austrian Science Fund FWF (grant I 4215) and the Slovenian Research Agency ARRS (grant N6-0113).

by Lowenstamm (2014), by answering the question what entry or entries for *ov* the Encyclopedia needs to contain.

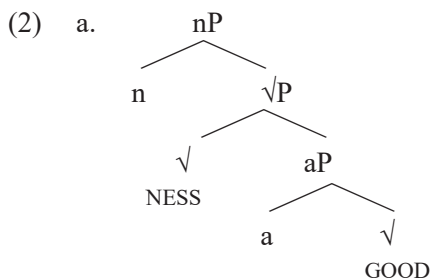
In what follows, we first give an overview of the account which treats affixes as roots in section 2. Then section 3 focuses on the morpheme *ov* in contexts which would be traditionally classified as derivational: derived verbs, possessive and kind adjectives and adjectives with the affix *ov(e)n*. Section 4 focuses on *ov* in nominal declensions. Section 5 is the conclusion.

## 2 DERIVATIONAL AFFIXES AS ROOTS

One of the postulates of Distributed Morphology is that roots do not carry information about the category of the word (Halle/Marantz 1993, 1994; Marantz 1996). Rather, the category of the word is determined by a categorial head. These heads can either have no phonological content (e.g. in *travel*, where the root  $\sqrt{\text{TRAVEL}}$  is combined with a mute categorizer *v* or *n*) or they can be phonologically realized, as is the case in most derivational affixes (e.g. in *traveler*). Recently, however, Lowenstamm (2014) showed that the assumption that derivational affixes are exponents of categorial heads runs into problems and instead proposed that derivational affixes are a subset of roots. Put differently, Lowenstamm (2014) proposes that derivational affixes are like roots in that they do not carry information about their category, but receive a category by being merged into the complement position of a categorial head. This move entails a separation between phonological/semantic content, on the one hand, and categorial heads, on the other, for both ‘traditional’ roots and affixes. Roots (including derivational affixes) have phonological content and/or meaning, while categorial heads are mute and have no semantic contribution. In this revised picture, affixes such as the English *ic* (which can be found in adjectives such as *atomic* below) are not analyzed as in (1a), but rather as in (1b), both of which taken from Lowenstamm (2014: 232, (6)).



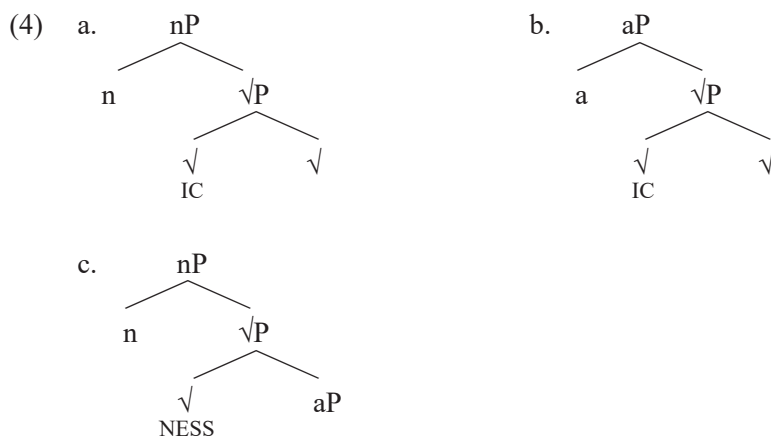
As (1b) already indicates, a crucial distinction between ‘traditional’ roots (e.g.  $\sqrt{\text{TRAVEL}}$ ) and affixes comes from the affixes’ requirement to take complements. Crucially, affixes can select roots, as *ic* in (1b), or categories, as is the case for *ness*, shown in (2) below.



One argument for the ‘affixes are roots’ treatment comes from the observation that certain derivational affixes, such as the English *ic*, surface under different categorial embeddings, as shown in (3). Note that in this respect affixes are no different from regular roots such as  $\sqrt{\text{WORK}}$  or  $\sqrt{\text{TRAVEL}}$ .

- (3) a. n: comic, academic, basics, sceptic  
 b. a: comic, academic, basic, atomic

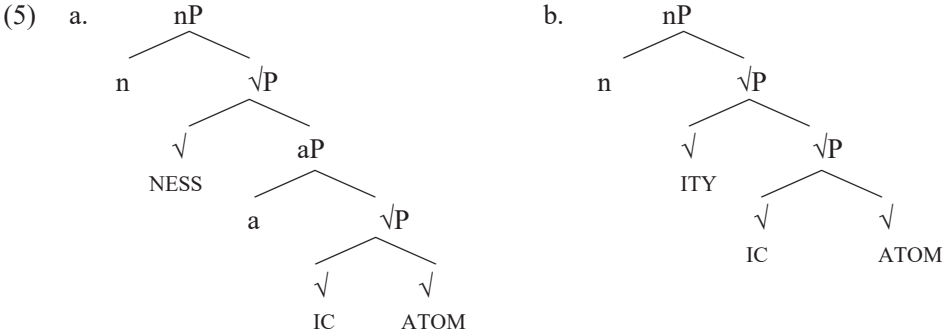
One consequence of treating derivational affixes as roots is the necessity of storing them as such, an issue not explicitly tackled by Lowenstamm. Simonović (2020) argues that the categorial embeddings of affixal roots need to be stored in the Encyclopedia. Storing the categorial embedding of roots is classically assumed in DM for ‘traditional’ roots (such as  $\sqrt{\text{CAT}}$ ), as summarized in the title of Marantz (1996) ‘*Cat* is a phrasal idiom’. Extending this to affixal roots, Simonović suggests that the categorial embedding of affixal roots should not be stored for each derived word, but once in an abstract schema, which also specifies the selectional behavior of the root. The Encyclopedia entries for the nominal and adjectival *ic*, and for *ness* would look as represented in (4).



These Encyclopedia entries raise the further question of their stored meaning. Simonović suggests that, unlike ‘traditional’ roots, affixal roots are potentially stored without any meaning. This is compatible with the observation that the meaning of affixal roots is often extremely vague or unpredictable (especially of those that appear under different categorial embeddings). For example, as Creemers et al. (2017: 75) note, “[affixes such as *ic*] have, among other interpretations, meanings such as ‘of,’ ‘relating to,’ ‘engaged in,’ and ‘connected with.’”

Assuming (nearly) meaningless roots then leads to one further question, i.e. why such roots do not appear in all environments without any limitations. The tentative answer is that their insertion generally leads to well-formed items, but some items are dispreferred if a less complex structure is available. However, as will be shown, certain (nearly) meaningless roots do indeed appear in a variety of contexts, including inflection. One example of such an affixal root is *ov*, which we will consider in section 3.

The second argument for the root analysis of affixes comes from English stress. While we will not go into the details of Lowenstamm’s account here, suffice it to say that under his approach the combination of affixes-as-roots analysis and phrasal spell-out can account for the prosodic contrast between stress-affecting affixes, (e.g. *ity* in *atomicity*) and stress-neutral ones (e.g. *ness* in *atomicness*). The relevant trees are shown in (5) below. The stress-affecting behavior is not essentially a property of the affix *ity*, but of the structure it appears in. The structures with multiple roots heading each other, termed “radical cores” by Lowenstamm, always constitute a single phase and therefore a single stress-assignment domain. Here Lowenstamm assumes that phonological rules re-apply with each new root phrase, but the same result can be obtained by having the rules apply to the whole radical core at once, on the first spell-out.

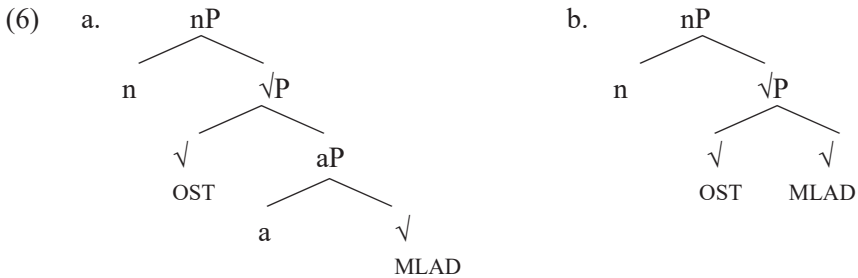


Crucially for what follows, Simonović (2020) extends this approach to Slovenian and shows how observations about word stress first made in Marvin (2003) can be accounted for if we treat affixes as roots. Marvin observes that there exists a prosodic contrast in (apparently) deadjectival *ost*-nominalizations, illustrated by minimal pairs

such as *mládost* ‘younghness’ (prosodically faithful to the adjective *mlád* ‘young’) and *mladóšt* ‘youth, young years’. Marvin’s analysis, which Simonović follows, is that *mládost* is a deadjectival nominalization (i.e. the more productive type, compositionally interpreted and prosodically faithful to the base adjective), while *mladóšt* is a root nominalization (i.e. the more rare, idiomatic type, characterized by stress-shifting). For Marvin, the stress-shifting behavior of *ost* is a consequence of a prosodic specification on the suffix, which is only realized phase-internally. Simonović dispenses with prosodic specifications on affixes altogether and proposes that the ‘idiomatic’ nominalization *mladóšt* contains a radical core (a combination of two roots) and that radical cores always receive the default stress pattern.

In Slovenian, the default stress is stem-final. Stem-final stress is also the most common prosodic pattern in the language, as established based on the stress pattern of the 3,000 most frequent nouns, verbs and adjectives. For each of these 9,000 words, we marked the stress pattern and annotated whether stress is stem-final. Items which can either have stem-final stress or another stress pattern were excluded from the count. In each of the three categories the stem-final stress pattern is by far the most common one, and a majority of words have this pattern, specifically, 63% of verbs, 70% of nouns and 73% of adjectives receive the stem-final stress.

While *mladóšt* has a default stress pattern, *mládost* is a deadjectival nominalization in which the root *ost* takes an adjective as its complement, which naturally leads to faithful prosody. The relevant trees are shown in (6).



To sum up the theoretical background presented in this section, we assume an approach under which derivational affixes can be treated as roots with potentially little or no meaning and according to which word stress is dependent on the structure of the word. In section 3, we show how this approach successfully captures the behavior of *ov* in the verbal and adjectival domain in Slovenian.

### 3 SLOVENIAN AFFIX *OV* IN THE DERIVATIONAL MORPHOLOGY

Before proceeding to the contexts in which *ov* can be found, we first need to note that *ov* surfaces either as *ov* or as *ev*, depending on the preceding consonant. This is a purely phonologically driven allomorphy, as illustrated in the possessive adjectives in (7).

- (7) a. Vid-ov      b. Rok-ov      c. Maj-ev  
       ‘Vid’s’      ‘Rok’s’      ‘Maj’s’

In traditional grammars, *ov* was in some instances taken to be a morpheme, for example in possessive adjectives as in examples such as (7) above, but it was primarily treated as a part of larger morphemes. For example, Toporišič (2000: 184) lists several affixes that are used to derive collective nouns. Among these we can find *je* (as in *cvet-je* ‘flowers’, related to *cvet* ‘flower’) but also *evje* and *ovje* (as in *borovničevje* ‘blueberry plants’, related to *borovnica* ‘blueberry’, and *cvetovje* ‘flowers’, related to *cvet* ‘flower’), *stvo* (as in *članstvo* ‘members’, related to *član* ‘member’) and *ovstvo* (as in *judovstvo* ‘Judaism’, related to *jud* ‘Jew’), while *ov* is not listed as a morpheme in this context. Notably, such treatment suggests that *stvo* as a morpheme has nothing in common with *ovstvo*. Marvin (2003) makes a similar point regarding morphemes *ec* and *je* in deverbal nominalizations.

In order to avoid the undesirable reduplication of affixes, we assume as a null hypothesis that all instances of *ov* are instances of the same derivational affix, which can also combine with other derivational affixes. Especially illustrative of this behavior are the pairs with and without *ov* which have slightly different meanings and can be found among denominal adjectives (8), collective nouns (9) and denominal nouns (10). In each of these contexts, it is unclear what the semantic contribution of *ov* is or what conditions its presence.

- (8) a. jezik-ov-en                      b. jezič-en  
       language-*ov-en*                  language-*en*  
       ‘related to language’          ‘related to tongue’
- (9) a. grm-ov-je                        b. sad-je  
       bush-*ov-je*                        fruit-*je*  
       ‘shrubbery’                        ‘fruit’
- (10) a. bank-ov-ec                      b. obraz-ec  
       bank-*ov-ec*                        face-*ec*  
       ‘banknote’                        ‘form’

In the above contexts, the affix *ov* is sandwiched between other categorized elements, therefore revealing no category with which it is associated. There are, however, instances of *ov* which would be analyzed as categorizers in classical Distributed Morphology. Such is the *ov* in possessive adjectives illustrated in (7). A similar analysis is plausible for denominal verbs in (11).

- (11) a. pot-ov-a-ti                      (cf. pot)      b. glas-ov-a-ti                      (cf. glas)  
       travel-*ov*-THEME-INF              vote-*ov*-THEME-INF  
       ‘to travel’                              ‘path, travel’      ‘to vote’                      ‘voice, vote’



There appear to be no clear cases in which *ov* functions as a nominalizer (but we will argue that such cases actually exist in the inflectional domain in Section 4).

In sum, the derivational uses of *ov* point towards *ov* being an extremely multifunctional affix, comparable to the English *ic*, i.e. a root that has little to no semantic contribution, which can appear in various categorial contexts.

As is clear from the examples above, the contexts in which *ov* shows up are extremely numerous and analyzing all of them would go beyond the scope of this article. We therefore made a representative selection of the derivational uses of *ov*. In 3.1, we turn to the verbal *ov*, which functions as a verbalizer and as an imperfectivizer. In 3.2 we offer an analysis of the adjectival *ov* in denominal adjectives. Finally, in 3.3 we turn to a case where *ov* itself reveals no category, as it is followed by the adjectivizing affix *-n*.

In each case our discussion will be guided by the question what the prosodic and semantic effects of *ov* are and what needs to be stored in the Encyclopedia in order to obtain these effects.

### 3.1 The affix *ov* in verbs

Before turning to the combinatorial possibilities of the verbal *ov*, it should be pointed out that *ov* in verbal contexts is characterized by allomorphy. The version *ov* shows up in the non-finite forms, where it is accompanied by the theme vowel *a*, (12a), whereas the version *u* shows up in the non-finite forms, where it is accompanied by the theme vowel *je* (12b).

- |  |   |
|--|---|
| (12) a. pot-ov-á-ti<br>travel- <i>ov</i> -THEME-INF<br>'to travel' | b. pot-ú-je-mo<br>travel- <i>u</i> -THEME-1SG.PRES<br>'we travel' |
|--|---|

As the examples in (12) show, the two allomorphs display different prosodic patterns: while in the non-finite forms the theme vowel is stressed, the stem-final vowel receives the stress in the present tense. As argued by Simonović (2020), these prosodic patterns are not a feature of the morpheme *ov*, but rather imposed by the theme vowels, since they are also attested with other roots (e.g. *or-á-ti* 'to plough', *ór-je-mo* 'we plough').

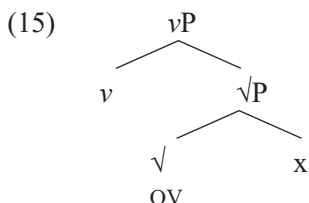
When it comes to the categories which the verbalizing *ov* selects and the effects it has, there seem to be two large classes. In most cases, *ov* shows up as an imperfectivizer, (13), which derives imperfective verbs (13a, c) from perfective ones (13b, d).

- |  |  |
|--|--|
| (13) a. kup-ov-ati<br>buy- <i>ov</i> -THEME-INF<br>'to buy.IMPERF'           | b. kup-i-ti<br>buy-THEME-INF<br>'to buy.PERF'                |
| c. pre-pis-ov-ati<br>over-write- <i>ov</i> -THEME-INF<br>'to copy' (imperf.) | d. pre-pis-a-ti<br>over-write-THEME-INF<br>'to copy' (perf.) |

In other cases *ov* shows up as a verbalizer, as illustrated in (14), taking bases of all other categories: nouns, (14a), adjectives, (14b), phrases, (14c) and roots which do not surface as independent words, (14d).

- |         |             |              |              |               |               |
|---------|-------------|--------------|--------------|---------------|---------------|
| (14) a. | pot-ov-a-ti | (cf. pot)    | b.           | modr-ov-a-ti  | (cf. moder)   |
|         | ‘to travel’ | ‘travel’     |              | ‘to theorize’ | ‘wise’        |
|         | c.          | vseb-ov-a-ti | (cf. v sebi) | d.            | spošt-ov-a-ti |
|         |             | ‘to contain’ | ‘in oneself’ |               | ‘to respect’  |

At first sight the imperfectivizing function seems to require a separate entry in the Encyclopedia. However, as proposed by Arsenijević (2018) for Serbo-Croatian, imperfectivizers can be analyzed as re-verbalizers, which reverse the value of the verbal aspect to its default, which is imperfective. If this is the case, then the verbalized *ov* is a good example of a universal selector, which can have any kind of element (phrase or root) in its complement. The Encyclopedia entry of the verbal *ov* is shown in (15), where *x* represents any element.



Before turning to the adjectivized *ov*, a remark is in order concerning the stress of *ov*-verbs. The stress pattern described above and imposed by the theme vowels is preferred for all *ov*-verbs by all Slovenian speakers, and a majority of the consulted speakers report not ever using any other stress pattern on these verbs. However, some speakers (from Upper Carniola, Lower Carniola and Ljubljana) allow the exceptional pattern with the stress on the syllable preceding *ov*. These speakers have realizations such as *pót-ov-a-ti* ‘to travel’ and *vér-ov-a-ti* ‘to believe’, next to more common *pot-ov-á-ti* and *ver-ov-á-ti*. We take this exceptional stress as evidence of the incorporation of an nP into the verb (in these cases of *pót* ‘travel’ and *vér-a* ‘faith’, respectively), which enables the preservation of the nominal stress. This is not unexpected given the cross-linguistic evidence that nominal lexical stress tends to be more strongly protected than that of verbs (Smith 2011). Assuming that the incorporation of nPs is the general way of capturing exceptional (i.e. not theme-controlled) stress in Slovenian verbs leads to the postulation of some nPs which do not surface independently. In our data set *var* and *napred* in the verbs *vár-ov-a-ti* ‘to guard’ and *napréd-ov-a-ti* ‘to make progress’ are instances of such non-attested nPs. However, while *var* and *napred* are not attested nouns, they are attested (with the expected stress pattern) inside adjectives *vár-en* ‘safe’ and *napréd-en* ‘progressive’, respectively.

### 3.2 The affix *ov* in possessive and kind adjectives

As already shown in the examples in (7), *ov* can be found in possessive adjectives, which are derived from either masculine or neuter animate nouns (but not feminine, which take the suffix *-in*). In (16) we provide further examples, together with bases, shown in the genitive singular form.

- |                                  |                                  |                              |
|----------------------------------|----------------------------------|------------------------------|
| (16) a. kralj-a<br>king.M.SG.GEN | b. vladark-e<br>ruler.F.SG.GEN   | c. deklet-a<br>girl.N.SG.GEN |
| a'. kralj-ev<br>'king's'         | b'. vladark-in/*-ov<br>'ruler's' | c'. deklet-ov<br>'girl's'    |

Kind adjectives formed with *ov*, on the other hand, are derived from nouns of either masculine, feminine or neuter gender. Nouns in these cases are interpreted as mass nouns and the adjectives generally have the interpretation 'made of x'.

- |                                  |                              |                               |
|----------------------------------|------------------------------|-------------------------------|
| (17) a. fižol-a<br>bean.M.SG.GEN | b. limon-e<br>lemon.F.SG.GEN | c. žvepl-a<br>sulfur.N.SG.GEN |
| a'. fižol-ov<br>'bean'           | b'. limon-ov<br>'lemon'      | c'. žvepl-ov<br>'sulfur'      |

Given the two groups of adjectives, which behave differently with respect to the gender of the noun they are derived from, the first option to be considered is that (i) possessive adjectives are derived from nPs (18a), and (ii) kind adjectives are derived from roots, (18b).

- |         |  |    |  |
|---------|--|----|--|
| (18) a. |  | b. |  |
|---------|--|----|--|

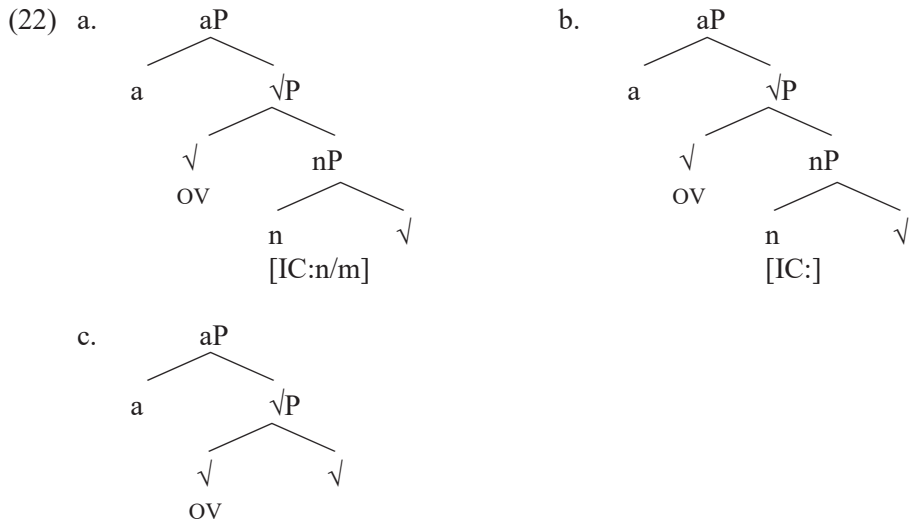
However, as we have seen in section 2 and as proposed in Simonović (2020), radical cores (i.e. roots directly adjacent to roots) trigger default stress, which in Slovenian is stem-final stress. This means that we would expect adjectives such as *fižolov*, (17b'), to be pronounced as *\*fižolóv* (to be more precise, *\*fižol[ów]*), which is not the case. In fact, the stress pattern of the two kind and possessive *ov*-adjectivizations is the same (and faithful to the stress of the nominal base).

- (19) a. *králj-év* 'king's'      b. *vladárk-in/\*-ov* 'ruler's'      c. *deklét-ov* 'girl's'
- (20) a. *fižól-ov* 'bean'      b. *limón-ov* 'lemon'      c. *žvépl-ov* 'sulfur'

Furthermore, kind adjectives always take nominal bases (and not bound roots or other categories), which again indicates that the structure in (18b) is not accurate. There is only a very small class of *ov*-adjectives which does have all the expected features of root adjectivizations (such as stem-final stress):

- (21) a. *kralj-év* 'royal'      (cf. *kralj* 'king')
- b. *njeg-[ǫ]v* 'his'      (cf. *nj-ega* 'him')
- c. *kak-[ǫ]v-ost* 'quality'      (implying the unattested adjective *kak-óv*)

This means that we need three distinct structures in order to capture the three groups of adjectives and two of them need to include an nP. We suggest that nPs can be selected either with the inflectional class/gender specified or without such a specification. Then, the structure for possessive adjectives (*králj-év*) incorporates an nP with a declension class specified on the n, see (22a). On the other hand, the structure for kind adjectives (*fižól-ov*) incorporates an nP *without* a declension class specified on the n, (22b), which explains the fact that kind adjectives can also be derived from feminine bases. Finally, adjectives like *kralj-év* are genuine root adjectivizations and are all stored with a specific meaning.



As for the meaning of these adjectives, we argue that the morpheme *ov* has no concrete semantic contribution, but that the possessive and kind meanings are a consequence of the structure. That is, we suggest that the possessive meaning might be the default meaning for an adjective derived from an animate noun. This seems especially plausible given possessive constructions in other languages where no overt possessive morphology is used, but rather simple adjacency. Such a language is Egyptian Arabic in which possessive phrases have the structure as in (23).

- (23) a. *kitaab Hasan*      b. *kitaab il-walad*  
       book Hasan            book the-boy  
       ‘Hasan’s book’        ‘the boy’s book’

Similarly, we suggest that the mass meaning of kind adjectives is essentially a default. This meaning is achieved by simple adjacency in a variety of languages, including English, as shown in (24).

- (24) a. *bean soup*                      b. *lemon juice*

As for the third group, listed in (21), their meaning is simply stored.

### 3.3 The affix *ov* with *n* in adjectives

While staying in the adjectival domain, we now focus on a context in which *ov* does not appear to be immediately headed by any category, as it is selected by another root. The selecting root in this case is  $\sqrt{N}$ , which is part of the most general adjectivizer in Slovenian (surfacing as *en* in the citation form). As mentioned in Section 2, *en*-adjectives are one of the contexts in which minimal pairs with and without  $\sqrt{ov}$  are attested, as illustrated in (25).

- (25) a. *jezik*                      b. *jezik-ov-en*                      c. *jezič-en*  
       ‘language, tongue’        ‘related to language’        ‘related to tongue’  
  
       a’.*plod*                      b’.*plod-ov-en*                      c’.*plod-en*  
       ‘fruit’                      ‘related to a/the fruit’        ‘fertile’

Not surprisingly, in the face of such data, traditional grammars list three different affixes used to derive adjectives that express ‘the relation to what the noun (from which the adjective is derived) expresses’: *oven* (as in *časoven* ‘temporal’ related to *čas* ‘time’), but also *en* (as in *čajen* ‘tea [adjective]’, related to *čaj* ‘tea’) and *ov* (*ogljikov* ‘carbon [adjective]’, related to *ogljik* ‘carbon’) (Toporišič 2000: 197–198).

Our analysis only assumes roots  $\sqrt{N}$  and  $\sqrt{ov}$ , which can be in the complement of an adjectival head, but can also combine with each other in more than one configuration, as shown by the prosody. Adjectives in *oven* display prosodic variation with a major and a minor pattern, comparable to the situation in the verbs in *ovati* in 3.1 and

adjectives in *ov* in 3.2. The major pattern is stem-final prosody, while a few items have either optional or obligatory preservation of the prosodic pattern of the base noun. This once again points in the direction of the exceptional preservation of nominal prosody under other categorial embeddings. The three possibilities are illustrated in (26), where all the adjectives are in the definite form because the citation form (e.g. *posloven*), which has no ending, contains an epenthetic shwa vowel, which obscures the stem-final stress position.

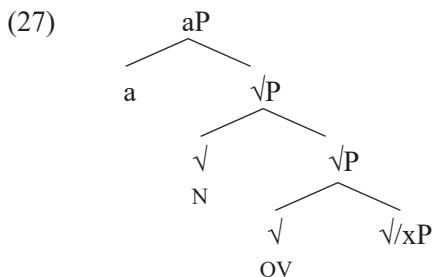
- (26) a. posl-[ǎ]v-n-i      b. dél-ov-n-i      c. blók-ov-n-i/blok-[ǎ]v-n-i  
       ‘business-related’    ‘work-related’    ‘bloc-related’

In order to establish the quantitative relations between the three patterns, we extracted the adjectives in *-oven* attested more than ten times in the Slovenian national corpus Gigafida. The search yielded 210 adjectives which were plausibly analyzable as containing *ov+n*. The relevant figures are shown in the Table 1 below.

Table 1: Stress on *ov+n*-adjectives

Stress pattern	Stem-final	On the base noun	Either stem-final or on the base noun
Number of adjectives	195	11	4

Given the stem-final stress on the majority of the *oven*-adjectives, we propose that the structure in (27) is the stored structure. Note that we remain agnostic as to which category *ov* selects (a root or an nP), as we do not have enough data to tease apart these two options. More generally, the well-formedness of structures in which a category appears below a radical core and their predicted stress pattern need to be addressed by further research.



The instances of *ov* in what is traditionally considered to be derivation show that *ov* can be analyzed as a single root which has no specific meaning, can select different categories or roots and can in turn appear embedded under a root such as  $\sqrt{N}$  or different categories. In the next section, we turn to *ov* in inflection.

#### 4 THE AFFIX *OV* IN THE NOMINAL DECLENSION

In inflection we can observe *ov* acting as the genitive case ending in dual and plural of the main masculine declension:

(28)		‘hill’	‘address’
	NOM.SG	hrib	naslov
	GEN.SG	hrib-a	naslov-a
	NOM.DU	hrib-a	naslov-a
	GEN.DU	hrib-ov	naslov-ov
	NOM.PL	hrib-i	naslov-i
	GEN.PL	hrib-ov	naslov-ov

Additionally, we can also observe *ov* in about 40 monosyllabic nouns, which take an *ov* augment in dual and plural (see Mirtič 2016 for a recent list of nouns taking the augment).

(29)		‘edge’
	NOM.SG	rob
	NOM.DU	rob-[ǎ]v-a
	NOM.PL	rob-[ǎ]v-i

The genitive and the augment *ov* display an interaction. In the plural dual/paradigms where there is an augment, the genitive form only contains one *ov*, as shown by the genitive dual/plural form *rob*-[ǎ]v in (30).

(30)		‘hill’	‘address’	‘edge’
	NOM.SG	hrib	nasl[ǎ]v	rób
	GEN.SG	hrib-a	nasl[ǎ]v-a	rób-a
	NOM.DU	hrib-a	nasl[ǎ]v-a	rob-[ǎ]v-a
	GEN.DU	hrib-ov	nasl[ǎ]v-ov	rob-[ǎ]v ( <sup>??</sup> rob-ov-ov)
	NOM.PL	hrib-i	nasl[ǎ]v-i	rob-[ǎ]v-i
	GEN.PL	hrib-ov	nasl[ǎ]v-ov	rob-[ǎ]v ( <sup>??</sup> rob-ov-ov)

Crucially, this is not due to haplology, as examples such as *nasl*[ǎ]v-*ov* ‘address.GEN.DU/PL’ show. The natural question is then which of the two *ov*s is pronounced in *rob*-[ǎ]v ‘edge.GEN.DU/PL’. The stress pattern provides a clear clue. The augment *ov* is always stressed, as can be seen from the augmented forms of *rob* in the table above. On the other hand, the genitive *ov* is never stressed, as can be seen from the forms *hrib*-*ov* ‘hill.GEN.DU/PL’ and *nasl*[ǎ]v-*ov* ‘address.GEN.DU/PL’ (the only exceptions being the few nouns in which the ending contains the only stressable syllable nucleus in the word, e.g. *ps*-a ‘dog.NOM.DU’, *ps*-i ‘dog.NOM.PL’, *ps*-[ǎ]v ‘dog.GEN.DU/PL’). The stress pattern in *rob*-[ǎ]v ‘edge.GEN.DU/PL’ is the same as in all the forms with the augment, but different

from all forms with just a case ending (e.g. the genitive singular *rób-a*), from which we can infer that it is the genitive *ov* that is deleted and hence that it is the augment *ov* that survives.

(31)		SINGULAR	DUAL	PLURAL
	NOMINATIVE	rób	rob-[ǰ]v-a	rob-[ǰ]v-i
	GENITIVE	rób-a	rob-[ǰ]v	rob-[ǰ]v
	DATIVE	rób-u	rob-[ǰ]v-oma	rob-[ǰ]v-om
	ACCUSATIVE	rób	rob-[ǰ]v-a	rob-[ǰ]v-e
	LOCATIVE	rób-u	rob-[ǰ]v-ih	rob-[ǰ]v-ih
	INSTRUMENTAL	rób-om	rob-[ǰ]v-oma	rob-[ǰ]v-i

The question is then how we can account for the two instances of *ov* in the nominal declension, as well as their interaction. Starting with the genitive *ov*, our proposal is that *ov* is the Elsewhere allomorph in the nominal paradigm and that its insertion is conditioned by phonological constraints. The Vocabulary Item for this item is as shown in (32).

(32)	/ov/ ↔ [ ]
------	------------

The assumption that there is no specialized genitive dual or plural ending in Slovenian is confirmed by the broader picture. The genitive dual and plural form is typically the form with no ending in both most common feminine (e.g. *ženska* ‘woman’) and most common neuter (*delo* ‘work’) declensions, but also in some nouns of the masculine declension class to which nouns such as *hrib* ‘hill’ belong. In addition to the ones that receive the augment *ov*, we find zero genitive plural/dual ending in *otrok* ‘child’, *konj* ‘horse’, *las* ‘hair’, *zob* ‘tooth’ etc. (Toporišič 2000: 283).<sup>2</sup> Below we show examples of plural paradigms with a zero ending in the genitive form.

(33)		‘teeth’	‘women’	‘works’
	NOMINATIVE	zob-je	žensk-e	del-a
	GENITIVE	zob	žensk	del
	DATIVE	zob-em	žensk-am	del-om
	ACCUSATIVE	zob-e	žensk-e	del-a
	LOCATIVE	zob-eh	žensk-ah	del-ih
	INSTRUMENTAL	zob-mi	žensk-ami	del-i

In sum, based on this lack of an overt genitive ending in two major feminine and neuter declension classes and in several exceptions in the masculine class, we can

<sup>2</sup> Additional evidence for this claim comes from plurale tantum nouns such as *možgan-i* ‘brain’, which can have a bare genitive plural *možgan* (but also *možgan-ov*).



assume that Slovenian has no specialized genitive dual/plural ending, and that what shows up in *gen.du/pl* is the Elsewhere allomorph.

Importantly, the insertion of the Elsewhere allomorph is restricted to very few cases, while being generally blocked. The competition between forms including the Elsewhere allomorph and those without it can be modeled in an Optimality Theory model. The issue of the relation between Lexical Insertion and phonology proper is a complex one (see Wolf 2013 for an overview). For the simplicity, we assume here that phonology (construed as an OT grammar) evaluates the candidates which are results of Lexical Insertion. Since Lexical Insertion is guided by the Subset principle, there is, as far as the Subset principle is concerned, no limit on inserting Elsewhere allomorphs. What filters those Elsewhere allomorphs out in most forms is a phonological constraint. One of the constraints that can be used to model this is the Optimality Theory's classic \*STRUCTURE (Zoll 1992; Prince/Smolensky 1993). This constraint militates against structure in general, always picking the candidate that has less structure. Given the Vocabulary Item in (32), any number of additions of *ov* are lexically sponsored, so they will never incur a violation of Faithfulness. However, each of them will incur a violation of \*STRUCTURE, which, for the purpose of this example we assume to assign a violation mark for each morpheme. In (34) this is illustrated using the tableau for the Instrumental Singular form of *hrib*.

(34)	<i>hrib + om</i>	FAITHFULNESS	*STRUCTURE
	a. ▣ <i>hribom</i>		**
	b. <i>hribovom</i>		***!
	c. <i>ovhribovom</i>		***!*
	d. <i>hribomovovov</i>		***!**
	e. <i>hrib</i>	*!	

The only situations in which the elsewhere allomorph can survive in Slovenian are cases in which its addition prevents a violation of another constraint. In this case, the relevant constraint is one militating against syncretism with the citation form.

- (35) Contrast-Citation (CONC): Incur a violation if an output form A is segmentally identical to the citation form of the lexeme that A belongs to. (Pertsova 2015)

Pertsova (2015) proposed this constraint based on Russian Genitive plural. Interestingly, Pertsova shows that, diachronically, Slavic genitive plural *ov* became generalized as a consequence of homonymy avoidance.

In Slovenian, the problem of syncretism of genitive dual/plural with nominative singular does not arise in declensions of the type *žensk-a* and *del-o* (as they have overt nominative singular endings). The situation in the type *hrib* is different, since leaving the genitive dual/plural without an overt ending would lead to a violation of CONC. This violation is avoided by allowing the Elsewhere allomorph in genitive dual/plural. The full picture obtained this way is that no big declension class in Slovenian violates CONC, as can be verified from (36).

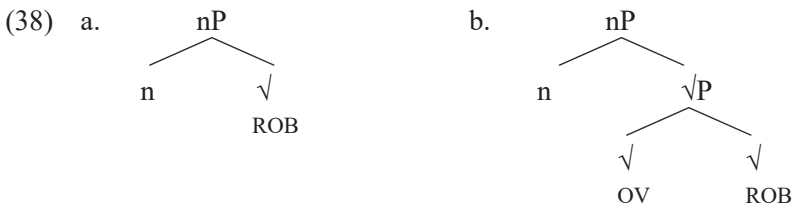
(36)		‘hill’	‘linden’	‘work’	‘thing’
	NOM.SG	hrib	lip-a	del-o	stvar
	GEN.SG	hrib-a	lip-e	del-a	stvar-i
	NOM.DU	hrīb-a	lip-i	deli	stvar-i
	GEN.DU	hrīb- <b>ov</b>	lip	del	stvar-i
	NOM.PL	hrīb-i	lip-e	del-a	stvar-i
	GEN.PL	hrīb- <b>ov</b>	lip	del	stvar-i

The tableau below shows the evaluation of the genitive dual/plural form of *hrib* form assuming the citation form *hrib*.

(37)	hrib	FAITHFULNESS	CONTRAST-CITATION	*STRUCTURE
	CF: hrib			
	a. hrib		*!	*
	b. $\blacksquare$ hribov			**
	c. hribovov			***!

The affix *ov* in the genitive dual/plural is therefore the Elsewhere morpheme used to avoid a CONC violation.

As for the always-stressed augment *ov*, we propose that nouns like *rob* ‘edge’ have two stored allomorphs and the phonological constraints choose between them (as in various phonological models of root allomorphy, e.g. Kager 2008). One allomorph is the root *rob* with its category, the second is the root *rob* in a root complex with the root  $\sqrt{ov}$ . As we have seen above, the augment *ov* is always stressed (*rob- $\acute{o}v$ -i*), which provides evidence for the root complex structure:



A crucial point for the evaluation below is that phonology does not have access to the internal structure of complex roots so it will consider the two stem allomorphs (*rób* ~ *rob $\acute{o}v$* ) as introducing the same amount of structure. As a result, the two allomorphs will be deployed to cover the two different positions in which no overt ending is inserted.

(39)	rób ~ robóv	FAITHFULNESS	CONTRAST-CITATION	*STRUCTURE
	CF: rob			
	a. rob		*!	*
	b. ▣ robóv			*
	c. róbov			**!
	c. robóvov			**!

Note that in the forms with an overt case ending, no preference can be established by the constraints. As a consequence, we assume that the forms which do have a preference (NOM.SG and GEN.DU/PL) impose their selected allomorphs upon the remainder of the number sub-paradigm through a uniformity constraint, which leads to the entire singular sub-paradigm taking the root allomorph *rob* and the entire dual and plural sub-paradigms taking the root allomorph *rob[ó]v*.

(40)		SINGULAR	DUAL	PLURAL
	NOMINATIVE	rób	rob- <i>ó</i> v-a	rob- <i>ó</i> v-i
	GENITIVE	rób-a	rob- <i>ó</i> v	rob- <i>ó</i> v
	DATIVE	rób-u	rob- <i>ó</i> v-oma	rob- <i>ó</i> v-om
	LOCATIVE	rób-u	rob- <i>ó</i> v-ih	rob- <i>ó</i> v-ih
	INSTRUMENTAL	rób-om	rob- <i>ó</i> v-oma	rob- <i>ó</i> v-i

In sum, in this section we have shown that  $\sqrt{ov}$  figures as a root with no specific meaning and as the Elsewhere allomorph in the nominal declension.

## 5 CONCLUSION

In this paper we have focused on the Slovenian affix *ov*, which can be found in nouns, adjectives and verbs. While these occurrences of *ov* appear *prima facie* to be several different affixes which only coincidentally have the same phonological form, we argue for a single multifunctional *ov*, which can appear both in inflection and derivation. In line with proposals within Distributed Morphology which claim that derivational affixes should be analyzed as roots, we have treated *ov* as a potentially meaningless root which can take as a complement other roots (thus forming a “radical core”) or phrases, resulting in different structures and consequently different stress patterns and meanings. In the nominal declension, *ov* acts as an Elsewhere allomorph, whose insertion is guided by an interplay of phonological and morphological constraints. Such a treatment of *ov* enables us both to give further support for the unification approach of affixes proposed in the literature and to further elaborate the affixes-as-roots model by addressing the question what the Encyclopedia entry or entries for *ov* need to contain.

Our consideration of the multifunctional morpheme *ov* in Slovenian is by no means intended as a definitive account (of this morpheme or of multifunctional morphemes in general). We therefore hope that further research will address our claims and predictions and broaden the data set, but also provide formal modeling of aspects we have

not explored here. One aspect of the account which has not been entirely formalized is the ability of the (semantically light) root to surface as the Elsewhere allomorph in the nominal paradigm. While semantically light roots are more plausibly expected to appear as Elsewhere allomorphs than roots with a fully specified meaning, we leave it to future research to account for the relevant mechanism. Slovenian seems a good starting point for such an account, as the same kind of parallelism is attested in the verbal domain (see Simonović, this volume).

### Primary sources

Gigafida. 24 February 2020. <http://www.gigafida.net/>.

### References

- ARSENIJEVIĆ, Boban (2018) “Serbo-Croatian suffix *-va* and the illusion of the imperfective specification.” Paper presented at the conference *Ambigo: Workshop on Ambiguity – Theory, Development, and Processing, University of Göttingen on 4–6 July 2018*. Göttingen: University of Göttingen.
- CREEMERS, Ava/Jan DON/Paula FENGER (2017) “Some affixes are roots, others are heads.” *Natural Language and Linguistic Theory* 36/1 45–84. <https://doi.org/10.1007/s11049-017-9372-1>
- HALLE, Morris/Alec MARANTZ (1993) “Distributed morphology and the pieces of inflection.” In: K. Hale/S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111–176.
- HALLE, Morris/Alec MARANTZ (1994) “Some key features of distributed morphology.” In: A. Carnie/H. Harley/T. Bures (eds), *MIT Working Papers in Linguistics* 21. Cambridge, MA: MITWPL, 275–288.
- KAGER, René (2008) “Lexical irregularity and the typology of contrast.” In: K. Hanson/S. Inkelas (eds), *The nature of the word: studies in honor of Paul Kiparsky*. Cambridge, MA: MIT Press, 397–432.
- LOWENSTAMM, Jean (2014) “Affixes as roots.” In: A. Alexiadou/H. Borer/F. Schäfer (eds), *The syntax of roots and the roots of syntax*. Oxford: Oxford University Press, 230–259.
- MARANTZ, Alec (1996) “‘Cat’ as a phrasal idiom.” Manuscript. Cambridge, MA: Massachusetts Institute of Technology.
- MARVIN, Tatjana (2003) *Topics in the stress and syntax of words*. Doctoral dissertation. Cambridge, MA: Massachusetts Institute of Technology.
- MIRTIČ, Tanja (2016) “Podaljševanje osnove z morfemom *-ov-* pri samostalnikih prve moške sklanjatve v sodobni knjižni slovenščini.” In: F. Marušič/P. Mišmaš/R. Žaucer (eds), *Škrabčevi dnevi 9. Zbornik prispevkov s simpozija 2015*. Nova Gorica: Založba Univerze v Novi Gorici, 16–33.
- NEVINS, Andrew (2015) “Free-Rides in Morphological Classification: The Case of Catalan Pre-Stressing Suffixes.” In: M. Lloret/C. Pons-Moll/E. Bosch-Roura (eds),

*Clàssics d'ahir i d'avui en la gramàtica del català*. Barcelona: Edicions Universitat Barcelona, 105–118.

PRINCE, Alan/Paul SMOLENSKY (1993) *Optimality Theory: Constraint Interaction in Generative Grammar*. Piscataway, NJ: Rutgers University Center for Cognitive Science.

SIMONOVIĆ, Marko (2020) “Derivational affixes as roots across categories.” [In press.]

SIMONOVIĆ, Marko/Boban Arsenijević (2020) “Syntax predicts prosody: Multi-purpose morphemes in Serbo-Croatian.” In: F. Marušič/P. Mišmaš/R. Žaucer (eds), *Advances in formal Slavic linguistics 2017*. Berlin: Language Science Press, 277–304.

SMITH, Jennifer (2011) “Category-specific effects.” In: M. van Oostendorp/C. J. Ewen/E. Hume/K. Rice (eds), *The Blackwell companion to phonology*. Oxford: Blackwell, 2439–2463.

TOPORIŠIČ, Jože (2000) *Slovenska slovnica*. Maribor: Obzorja.

WOLF, Matthew (2013) “Lexical insertion occurs in the phonological component.” In: E. Bonet/M. Lloret/J. Mascaró (eds), *Understanding allomorphy: perspectives from Optimality Theory*. London: Equinox Publishing, 361–407.

ZOLL, Cheryl (1992) “When syllables collide: a theory of alternating quantity.” Manuscript. Brandeis University.

#### Abstract

#### √OV IS IN THE AIR:

#### THE EXTREME MULTIFUNCTIONALITY OF THE SLOVENIAN AFFIX *OV*

In this paper we consider several instances of the Slovenian affix *ov*, which surfaces in many, apparently unrelated contexts. Here we focus on (i) *ov* in verbs, where it can act as an imperfectivizer or a verbalizer, (ii) *ov* found in possessive adjectives and kind adjectives derived from nouns, (iii) *ov* which precedes the adjectiviser (*e*)*n* in denominal adjectives, and (iv) *ov* in nominal declension (acting as a genitive case ending in dual and plural or as a dual/plural augment). Building on the observation that certain affixes function either as inflectional or as derivational (see Simonović and Arsenijević 2020), and working within a Distributed Morphology approach which postulates that derivational affixes should be analyzed as roots (e.g. Lowenstamm 2014), we argue for a single multifunctional *ov*. This *ov* is a potentially meaningless root that can take as a complement other roots (thus forming a “radical core”) or phrases, resulting in different structures and consequently different stress patterns and meanings, but can also act as an Elsewhere allomorph, whose insertion is guided by an interplay of phonological and morphological constraints.

**Keywords:** morphology, Distributed Morphology, Slovenian, multifunctional affix, roots

Povzetek  
EN ALI VEČ MORFEM- $\sqrt{OV}$ :  
IZJEMNA VEČFUNKCIJSKOST SLOVENSKEGA MORFEMA *OV*

V prispevku obravnavamo več funkcij slovenskega morfema *ov*, ki se pojavlja v številnih, med seboj navidezno nepovezanih okoljih. Osredotočamo se na (i) *ov* v glagolih, v katerih določa nedovršnost ali samo besedno vrsto, (ii) *ov*, ki ga najdemo v svojilnih in vrstnih pridevniki, izpeljanih iz samostalnikov, (iii) *ov*, ki se v izsamostalniških pridevniki pojavlja pred pridevniškim morfemom (*e*)*n*, in (iv) *ov* v samostalniških sklanjatvah, kjer služi kot rodilniška končnica v dvojini in množini ali za podaljšanje osnove. Na osnovi trditve, da so nekateri morfemi bodisi oblikotvorni bodisi besedotvorni (gl. Simonović in Arsenijević 2020), in v skladu s pristopom, ki v okviru t. i. razpršene morfologije predlaga, da je treba besedotvorne morfeme analizirati kot korene (npr. Lowenstamm 2014), trdimo, da obstaja v slovenščini en sam večfunkcijski morfem *ov*. Gre za potencialno brezpomenski koren, ki lahko za svoja dopolnila izbira druge korene (in tako tvori korenski skupek) ali zveze, kar vodi v različne strukture in se posledično odraža v različnih naglasnih vzorcih in pomenih. Poleg tega ima lahko morfem *ov* tudi vlogo zapolnjevalnega alomorfa, katerega rabo določajo tako fonološke kot morfološke omejitve.

**Ključne besede:** morfologija, razpršena morfologija, slovenščina, večfunkcijski morfem, koreni



## CATEGORIES, ROOT COMPLEXES AND DEFAULT STRESS: SLOVENIAN NOMINALIZATIONS REVISITED<sup>1</sup>

### 1 INTRODUCTION

Standard Slovenian is a language with lexical prosody in nouns, verbs and adjectives (the default prosodic pattern assigned when no lexical prosody is present or available is discussed in section 3.1). Further, derivational morphemes which behave as nominalizers, verbalizers and adjectivizers can have stress-shifting effects.

The interaction between lexical stress and the structure of words in Slovenian is discussed in Marvin (2003), where a detailed analysis of stress assignment in Slovenian nominalizations is presented. The nominalizations analysed by Marvin fall into two broad categories in terms of prosody: the nominalizing affix either leaves the prosodic pattern of the base unaffected or it causes the word stress to shift to a new position. In some cases, the same nominalizing affix can appear in both prosodic types, giving rise to minimal pairs that differ only in stress. It is exactly such cases that make a structural analysis of stress assignment all the more necessary, since the problem cannot be solved through lexical specification on either the base or the affix entailing multiple suspiciously homonymous affixes belonging to the same category.

The prosodic dichotomy between affixes is illustrated in (1). The nominalizations in (1a) inherit the prosodic pattern of the base, but those in (1b) have a different prosodic pattern from the base.

#### (1) Deadjectival and deverbal nominalizations in Slovenian

##### a. Stress-preserving nominalizations

Nominalization	Base	Nominalization	Base
stár-ost	stár	anketírán-ec	anketírán
old-ost		interview.PASS.PTCP-C	interview.PASS.PTCP
‘oldness’	‘old’	‘interviewee’	‘interviewed’

##### b. Stress-shifting nominalizations

Nominalization	Base	Nominalization	Base
star-óst	stár	napadál-ec	napádal
star-ost		attack.LPTCP-C	attack.LPTCP
‘old age’	‘old’	‘attacker’	‘attacked’

\* rkiema@gmail.com

1 I acknowledge the financial support by the Austrian Science Fund FWF (grant I 4215) and the Slovenian Research Agency ARRS (grant N6-0113).

Note that only the deadjectival nominalizations (*stár-ost* and *star-óst*) form a true minimal pair, as the deverbal nominalizations involve different participles. This is why Marvin’s analysis will first be summarised based on the pair *stár-ost* vs. *star-óst*. The crux of the analysis is that the stress pattern of complex words is negotiated as a function of Phasal Spell-out. An affix can only influence the prosody of the root if it is in the same phase with that root. This means that stress-preserving nominalizations contain several phases and the potentially stress-affecting affix is merged in a different phase from the root, i.e., it arrives ‘too late’ to change the spell-out of the root. For the pair *stár-ost* vs. *star-óst*, this means that the former is a truly deadjectival nominalization containing an adjectival and a nominal head, whereas the latter is a root nominalization (with no adjectival head). As illustrated in the trees in (2), in both *stár-ost* and *star-óst* the only two pieces of structure that have phonological content are the root  $\sqrt{\text{STAR}}$  and the nominalizer *-ost*. However, while in the deadjectival nominalization *stár-ost* they are separated by a phase boundary (aP being a phase), in the root nominalization *star-óst* they are adjacent. Assuming that *-ost* has underlying stress and is able to shift the stress from the adjacent head yields the right result. As Marvin points out, the difference between the two structures plays a role not only at the PF, but also at the LF interface. Indeed, in such pairs, the true deadjectival nominalizations tend to have a more transparent interpretation: *stár-ost* means the property of being *star* ‘old’, whereas the root nominalizations tend to have rather idiomatic meanings.

(2) *stár-ost* vs. *star-óst*: The analysis in terms of Marvin (2003)



As far as the deverbal nominalizations in (1) are concerned, it has already been noted that they do not form a perfect minimal pair, as they incorporate different participles. The logic of their analysis is therefore somewhat different, as the structure of the nominalization depends on the structure made available by the participle. As Marvin convincingly shows, only the passive participle has a truly adjectival structure, whereas the *l*-participle does not introduce a categorial head. This adjectival vs. non-adjectival contrast leads to a dichotomy similar to that between *stár-ost* and *star-óst* discussed above: the passive-participle nominalizations contain an adjectival phrase and therefore the nominalizer cannot be stress-shifting, whereas in the case of the *l*-participle nominalizations all pieces which have phonological content are in the same phase and the nominalizer can impose its own prosodic pattern.



One important assumption necessary for this account to work is that when categorial affixes and roots have conflicting prosodic requirements, the affixes win. This is not explicitly spelled out in Marvin (2003). However, since categorizers are heads, any model that assumes the phonological privilege of syntactic heads predicts this result. Such models have indeed been proposed, see Revithiadou (1999) for a cross-linguistically grounded proposal.

## 2 PREDICTIONS OF THE MODEL AND UNEXPECTED GAPS

Marvin’s analysis makes several predictions concerning the prosody of nominalizations. In this section I will consider Slovenian data against the background of these predictions and establish a list of empirical issues which need to be addressed when revisiting Marvin’s model.

Reasoning from the type of base, the predictions are as follows:

- A1) If the base is a root, we expect differences between stress-shifting and stress-preserving nominalizers (reflecting the lexical specification of the affix).
- A2) If the base is a root accompanied by a non-categorial affix (as with *l*-participle bases), the situation should be as with roots, i.e. we expect differences between stress-shifting and stress-preserving nominalizers (reflecting the lexical specification of the affix).
- A3) If the base has a category, we should always observe the stress-preserving behaviour of the nominalizers (as the lexical specification of the affix cannot be enforced across the phase boundary).

Some of the predictions above are easier to test than others. A1 is extremely hard to test because a root cannot easily be distinguished from a combination of a root and a phonologically mute categorial head. On the other hand, A2 can be tested using derivations which incorporate the *l*-participle. Here, the results are quite unexpected: all affixes, nominal and adjectival, which combine with the *l*-participle cause a stress shift. This is illustrated in (3) using the family of words which incorporate the *l*-participle *méril* ‘measured’.

(3) Derivations incorporating *méril* ‘measured’

Nouns:	<i>meríl-ec</i> measure.LPTCP- <i>c</i> ‘timekeeper’	<i>meríl-ka</i> measure.LPTCP- <i>ka</i> ‘timekeeper’ (fem.)	<i>meríl-o</i> measure.LPTCP- <i>o</i> ‘measure, scale’
Adjectives:	<i>meríl-n-a</i> measure.LPTCP- <i>n</i> -FEM ‘measuring’ (fem.)	<i>meríl-sk-a</i> measure.LPTCP- <i>sk</i> -FEM ‘related to measuring’	

The lack of stress preserving derivations with *l*-participles is the first empirical issue (EI1) which I seek to account for in this article by building on Marvin’s model.

E11: L-participle derivations are always stress-shifting.

A3 can be tested by isolating those bases which have an overt categorial head. If we remain focused on deadjectival *ost*-nominalizations and participle nominalizations, we expect the following:

- a) Derived adjectives (i.e. those adjectives which contain an overt adjectivizer) should never derive *ost*-nominalization with the stressed *ost*.
- b) Passive participles should never derive nominalizations with a stress-shift.

In order to test the first prediction, I extracted all *ost*-nominalizations which had more than 300 attestations in the Slovenian national corpus Gigafida. I annotated the 1,231 items obtained for their stress pattern (whether *-ost* is stressed) and the visibility of adjectivizing morphology (whether the base contains a common adjectivizing suffix).

Table 1: Stress and the visibility of the adjectivizer in *ost*-nominalization

Stress	Stress not on <i>-ost</i>		Stress on <i>-ost</i>	
	Adjectivizer visible	Adjectivizer not visible	Adjectivizer visible	Adjectivizer not visible
Example	oséb-n-ost person-adj-nom 'personality'	pogóšt-ost frequent-nom 'frequency'	last-n-óst property-adj-nom 'feature'	nor-óst insane-nom 'insanity'
Number	1079	126	4	22
Percentage (within the stress type)	90%	10%	15%	85%

The first observation to be made based on the above data is that the relative frequencies observed fit the model quite well, which means that what we need to handle are essentially exceptions. The only *ost*-nominalizations which contradict the model are the four nominalizations with a stressed affix *-ost* which cannot be analysed as root nominalizations since they contain clearly derived adjectives. Beside *last-n-óst* 'feature', these are *skriv-n-óst* 'secret', *dolž-n-óst* 'obligation' and *uč-en-óst* 'erudition'. While both their form and their meaning clearly point in the direction of idiomatized items, it is still an issue how we represent such items, as introducing new roots ( $\sqrt{\text{LASTN}}$ ,  $\sqrt{\text{SKRIVN}}$  etc.) just to fit the data seems a counter-intuitive move.

E12: There are (clearly exceptional) items in which the nominalizer *-ost* is stressed, although there is a visible adjectival head under it.

The nominalization *uč-en-óst* 'erudition' actually contains a passive participle *uč[é]n* 'learned' and is therefore an exception relevant for the next prediction we are

considering: passive participles should never form nominalizations with a stress-shift. Just like with the previous set, we are again dealing with exceptions, which require a representational solution. The examples of stress-shifting nominalizations from passive participle are very few, but they are attested in every type of nominalization. In the examples below, I use the nominalizations of the passive participle *mérjen* ‘measured’ in order to illustrate the regular pattern and the full list of exceptions I have identified (and which were confirmed as familiar to more than one third of the speakers I consulted).

(4) *c*-nominalizations from passive participles

a. Regular pattern (faithful prosody)		b. Exceptional pattern (shifted prosody)	
Base	Nominalization	Base	Nominalization
mérjen	mérjen-ec	múčēn	muč[é]n-ec
‘measured’	‘measured one’	‘tortured’	‘marthyr’

(5) *ka*-nominalizations from passive participles

a. Regular pattern (faithful prosody)		b. Exceptional pattern (shifted prosody)	
Base	Nominalization	Base	Nominalization
mérjen	mérjen-ka	múčēn	muč[é]n-ka
‘measured’	‘measured one’ (fem.)	‘tortured’	‘marthyr’ (fem.)

(6) *je*-nominalizations from passive participles

a. Regular pattern (faithful prosody)		b. Exceptional pattern (shifted prosody)	
Base	Nominalization	Base	Nominalization
mérjen	mérjen-je	míšljēn	mišlj[é]n-je
‘measured’	‘measuring’	‘thought’	‘opinion, thinking’
		vprašān	vprašān-je
		‘asked’	‘question’

This list of frequent exceptions leads us to the next empirical issue.

*E13: There are (clearly exceptional) items in which the nominalization of a passive participle is stress shifting.*

Another issue that arises from these data is the exceptional stress-shifting behaviour of many different affixes. In the examples above, beside *-ost*, these are, *-c*, *-ka* and *-je*. Especially the last one seems to behave as a stress shifter only in the two examples quoted above (plus potentially several words with a variable stress pattern). This makes it considerably less plausible to specify the stress-shifting behaviour in their lexical entry. Therefore, we can identify the last empirical issue.

*E14: Several affixes (e.g. -ost, -c, -ka and -je) act as stress-shifters in very few items.*

### 3 DESIDERATA FOR AN IMPROVED MODEL AND A NEW PROPOSAL

If the model of stress assignment in Slovenian discussed so far is to be improved, it is crucial not to lose its many virtues. The correlation between root-selecting behaviour and stress-shifting effects on the one hand, and adjective-selecting behaviour and stress-neutrality on the other is a robust fact confirmed by corpus data. I used the same methodology as for *ost*-nominalizations above on two more nominalizers the stress-shifting *-ota* (e.g. in *lep-ota* ‘beauty’) and stress-neutral *-ež* (e.g. in *perverz-n-ež* ‘pervert’). *Ota*-nominalizations show no clear examples of derived bases (based on 23 items), whereas *ež*-nominalizations have derived bases in 94% of the cases (based on 48 items).

There is, however, an overgeneration problem. The original model predicts the existence of root-selecting stress-neutral affixes. Recall that in this model the prosody of affixes was lexically specified and, while category-selecting affixes are necessarily stress-neutral, root-selecting affixes can be either stress-shifting or stress-neutral. Since I have not identified any stress-neutral nominalizers which generally select roots, I will present a proposal which excludes such a structure.

My proposal dispenses with lexical prosody on derivational affixes altogether. Instead, all the prosodic effects follow from the structural position in which the affix appears. Before I define the relevant positions, I first take a closer look at the prosodic effect which I have referred to as stress-shifting in the presentation of the data so far.

#### 3.1 Stem-Final Stress as the Default in Slovenian

All the data in which we found stress-shifting behaviour so far share the same stress pattern: stress is always stem-final, i.e. falling on the syllable preceding the inflectional ending. In some of the examples above, the stem-final position of stress was somewhat obscured in the citation form due to an epenthetic process in forms without overt inflectional morphology (e.g. in *muč[é]n[ə]c*, where the schwa is epenthetic and absent from all other forms, which have overt inflection). In (7) I provide the dual forms of examples of all the classes considered as stress-shifting in the text so far, since the dual ending always has an overt exponent. Note that in (7) the hyphen only separates the stem from the inflectional ending.

- (7) a. *ost*-nominalizations  
           staróst-i                      učénóst-i  
           ‘old ages’ (dual)    ‘eruditions’ (dual)
- b. derivations incorporating the *l*-participle *méril* ‘measured’  
           Nouns:  
           merílc-a                      merílk-i                      méríl-i  
           ‘timekeepers’ (dual)    ‘timekeepers’ (fem.) (dual)    ‘measures, scales’ (du.)
- Adjectives: meríln-i                      merílsk-i  
   ‘measuring’ (fem.) (dual)    ‘related to measuring’ (fem.) (dual)

c. nominalizations incorporating passive participles

muč[é]nc-a	muč[é]nk-i
‘martyrs’ (dual)	‘marthyr’s’ (fem.) (dual)
mišlj[é]nj-i	vprašanj-i
‘opinions’ (dual)	‘questions’ (dual)

d. *ota*-nominalizations

lep[ó]t-i	groz[ó]t-i
‘beauties’ (dual)	‘horrors’ (dual)

As the reader can testify, in all the examples above the stress is stem-final. This is by far the most common stress pattern in Slovenian and I argue that this is the Slovenian default prosodic pattern. Essentially, this means that stem-final stress is assigned to all words in which lexical prosody is not present or not available.

In order to substantiate the claim that stem-final stress is the most common one in Slovenian, Simonović and Mišmaš (this volume) isolated the 3,000 most frequent nouns, verbs and adjectives and marked the stress pattern of each category. Items that can either have the stem-final stress or another stress pattern were excluded from the count. In each of the categories the stem-final stress pattern is by far the most common one and a majority of words has this pattern. The stem-final stress pattern was encountered in 63% of the verbs, 70% of the nouns and 73% of the adjectives.

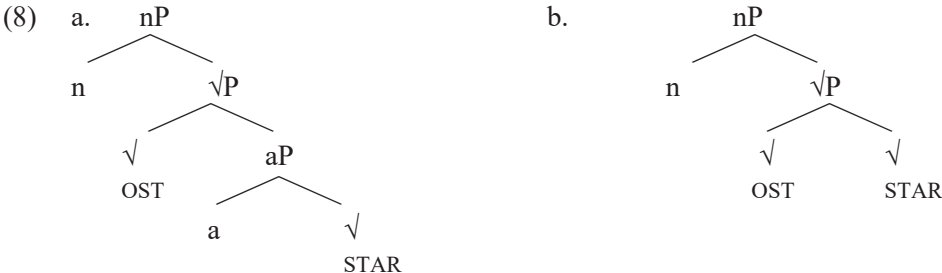
### 3.2 Derivational Affixes as Roots

Now that I have identified the stress pattern common to all the stress-shifting environments, I turn to the task of identifying what these environments have in common and considering why this should lead to the assignment of the default stress pattern. As it stands now, this class of environments looks quite heterogeneous, encompassing various nominalising affixes (*-ota*, *-ost*, *-c*, *-ka*, *-je*) with an environment which can only be described as a combination of the *l*-participle and any suffix which is added to it.

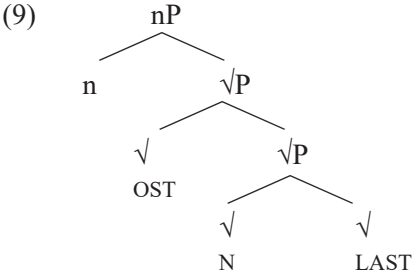
In a sense, following the intuition from Marvin’s proposal that stress-shifting is a phenomenon related to the root domain, I propose that what triggers the default prosody in Slovenian is indeed a root phenomenon, but of a kind that was not available in the theory used so far. Following Lowenstamm (2014), I assume that derivational affixes are transitive roots, some of which select categories, whereas others select other roots (see Simonović 2020) and Simonović and Mišmaš (this volume) for applications of this model to Slovenian). Lowenstamm’s model, originally proposed to resolve several issues in English nominalizations, predicts the existence of complex root phrases with no intervening functional structure. These structures are termed radical cores by Lowenstamm. My central assumption concerning the prosody of Slovenian radical cores is that they always receive default prosody. The idea is that the regular algorithm that computes compositional prosody requires the presence of functional projections that can decide which morphemes win in the event of conflict (much as in Revithiadou 1999, where heads win). Due to the disruption of the usual asymmetry caused by the

lack of functional structure, the prosody in radical cores cannot be computed compositionally and therefore the assignment of default stress ensues.

We can now revisit our first nominalization examples *stárost* and *staróst* and apply the new model to them. My analysis of *stárost* is basically still that of Marvin (2003) in that *stárost* is a deadjectival nominalization. The technical innovation illustrated in (8a) is that *-ost* is now a root that selects an adjective as its complement, which naturally leads to preserving the prosodic pattern of the base. The root nominalization *staróst* is still a root nominalization, but it now contains two roots,  $\sqrt{\text{STAR}}$  and  $\sqrt{\text{OST}}$ , which form a radical core, triggering default stress.



Note that, unlike Marvin’s model, the proposed model can represent the exceptional *ost*-nominalizations with an overt derivational affix preceding *-ost*, such as *last-n-óst* ‘feature’ (thereby resolving the Empirical Issue 2 identified in Section 2). In the four items like *last-n-óst*, *n* will be part of the radical core, as shown in (9).



While such items are representable, they are also expected to be rare, as they require a specific Encyclopedia entry. This prediction is borne out, as we only found four such items.

**4 CONSEQUENCES**

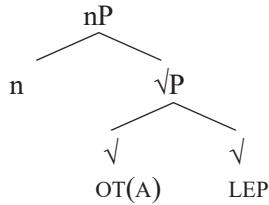
With the new proposed model in place and having applied it to the first portion of the data, I now turn to the more general picture.

The stress-shifting affixes we have seen so far fall into three different categories when it comes to their selectional properties:

- a) Affixal roots like *-ota* always select other roots so it seems safe to assume that they have the selectional feature [ $\sqrt{P}$ ].
- b) Affixal roots like *-ost*, *-c*, *-ka*, and *-je* select roots only in a subset of their uses, which can be seen as exceptions. These roots are then generally category selecting, having the feature [ $xP$ ], but they can be stored as part of radical cores as well.
- c) Combinations of the ending *-l* and other affixes generally shift stress.

Roots like *-ota* are entirely unproblematic and their stress pattern follows from the structure imposed by their selectional properties, as illustrated in (10).

(10)



The account for the other two types needs to be rather more elaborate, as there is more variation and interaction between several affixes. I address the exceptional formation of radical cores in 4.1. and the interactions of  $\sqrt{L}$  with other affixes in 4.2.

#### 4.1 Exceptional Radical Cores

The exceptional radical cores are readily analysed as cases of storage of a simplified structure, which contains just the bare minimum of functional projections, i.e. the top-most categorizing head, without which no well-formed word can exist.

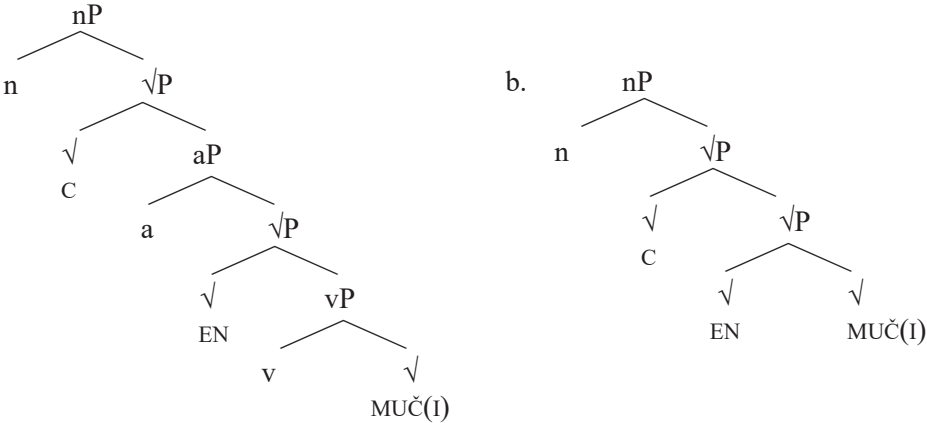
Note that I am now considering the *ost*-nominalizations with stressed *ost* (e.g. *staróst*) as exceptions. This is justified by the quantitative data presented in Table 1, which show that less than 3% of all *ost*-nominalization have the shifted stress pattern.

While in examples of the type *staróst* only one functional projection was removed from the structure, other examples involve several removed projections. In order to illustrate this, I quote a minimal pair that some of the speakers have involving two *c*-nominalizations of the passive participle *múčen* ‘tortured’.

- (11) Two possible *c*-nominalizations of *múčen* ‘tortured’
- | Base         | Nominalization 1          | Nominalization 2  |
|--------------|---------------------------|-------------------|
| <i>múčen</i> | <i>múčenc-a</i>           | <i>muč[é]nc-a</i> |
| ‘tortured’   | ‘tortured persons’ (dual) | ‘martyrs’ (dual)  |

Based both on the stress pattern and on meaning, it is clear that Nominalization 1 contains more functional structure than Nominalization 2. Nominalization 1 therefore also has more structure in common with the passive participle than Nominalization 2. The relevant trees for *múčēnec* and *muč[é]nec*, respectively are provided in (12).

(12) *múčēnec* ‘tortured persons’ vs *muč[é]nec* ‘marthyr’



Note that the proposed analysis is clearly distinct from the more traditional analysis which would simply state that the *muč[é]nec* is stored as a single item. Crucially, there is no new root  $\sqrt{\text{MÚČENC}}$  and the two items still contain the same roots, which are connected in a different structure.

This subsection addressed the Empirical Issues 3 and 4 from Section 2 by providing a model which can handle the exceptions, that is items exceptionally lacking functional layers present in more compositionally built related words. In the following subsection I address Empirical Issue 1, which concerns the stress-shifting behavior of the derivations that contain the *l*-participle.

**4.2 *L*-participle,  $\sqrt{L}$ , or Both?**

Turning now to the nominalizations which appear to be derived from *l*-participles, I can already formalize what was formulated as Empirical Issue 1 in Section 2: the fact that “*L*-participle nominalizations are always stress-shifting.” All that needs to be specified is that  $\sqrt{L}$  is a root-selector and the result of  $\sqrt{L}$  always triggering default stress is achieved. Now we can go through the contexts in which  $\sqrt{L}$  appears and see how plausible this assumption is.

While a combination of  $\sqrt{L}$  with other affixes is very common, there actually is a class of nominalizations which show no other (overt) derivational affix on top of  $\sqrt{L}$ .



One example quoted in (3) above was *merilo* ‘measure, scale’. Nominalizations of this type usually have the structure verbal root + theme vowel + *l* + *o* (*o* being a case ending). Marvin (2003: 108) notes that the theme vowel in some cases is not the same as the one attested in the verb. Furthermore, there are cases in which the root is not verbal at all. All three types are illustrated in (13).

(13) Three types of *lo*-nominalizations

a. verbal root + theme vowel attested with the same root + *lo*

Nominalization	Related <i>l</i> -participle
<i>merílo</i>	<i>méril</i>
‘measure scale’	‘measured’
<i>pisálo</i>	<i>písal</i>
‘writing device’	‘written’

b. verbal root + theme vowel not attested with the same root + *lo*

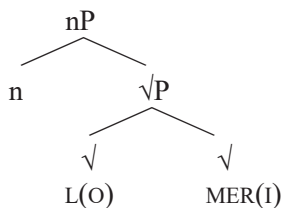
Nominalization	Related <i>l</i> -participle
<i>barvílo</i>	<i>bárval</i>
‘pigment’	‘coloured’
<i>tolkálo</i>	<i>tólkel</i>
‘percussion’	‘banged’

c. non-verbal root + theme vowel + *lo*

Nominal	Related <i>l</i> -participle	Related noun
<i>glasbíl-o</i>	/	<i>glásb-a</i>
‘musical instrument’		‘music’
<i>zrcál-o</i>	/	/
‘mirror’		

This variety of patterns attested in *lo*-nominalizations suggests a structure which is not necessarily deverbal, although it does seem to host roots which also show up in verbs more often than others. The analysis I propose is then that in (14). The inflectional ending *-o* is shown between brackets, which is a clear oversimplification: the *-o* is an exponent of a higher functional head (presumably *K*, the case head) which is not visible in the tree built so far.

(14)



I remain agnostic as to whether the root  $\sqrt{L}$  is the same as the morpheme at the end of *l*-participles, which Marvin (2003: 90) convincingly shows to be the Elsewhere allomorph in the verbal inflection. If this is the case, the picture is very well comparable to that in the nominal domain found by Simonović and Mišmaš (this volume) and it would necessitate a more general reconsideration of the distinction between the Encyclopedia and the Vocabulary. Note, however, that the lexical prosody of *l*-participles (e.g. *méri-l* ‘measured’ versus *mori-l* ‘murdered’) indicates that *l*-participles contain a vP under the position where the Elsewhere allomorph is inserted, so that there is no radical core in these forms and no default prosody emerges.

The most prominent feature of the root  $\sqrt{L}$  beyond *lo*-nominalizations is that it is very often selected by other roots, which then leads to whole families of related words as the one shown in (3) and repeated here in (15).

- (15) Derivations containing *méri+l* related to *mériti* ‘measure’
- |             |                   |                        |                  |
|-------------|-------------------|------------------------|------------------|
| Nouns:      | meríl-ec          | meríl-ka               | meríl-o          |
|             | ‘timekeeper’      | ‘timekeeper’ (fem)     | ‘measure, scale’ |
| Adjectives: | meríl-n-a         | meríl-sk-i             |                  |
|             | ‘measuring’ (fem) | ‘related to measuring’ |                  |

Especially productive are the agentive nominalizations in which  $\sqrt{L}$  is selected by  $\sqrt{C}$  and  $\sqrt{K(A)}$ : out of 3000 most frequent verbs in the corpus Gigafida, more than 500 have nominalizations in  $\sqrt{L}+\sqrt{C}$  (*-lec*) or  $\sqrt{L}+\sqrt{K(A)}$  (*-lka*), making this the most productive pattern for agent nominalizations in Slovenian. This being so, it may appear more plausible to assume a verbal structure under  $\sqrt{L}$ . While this may be an option worth exploring, there are several good reasons for sticking with Lowenstamm’s initial model in which no categories are allowed under radical cores. First, nominalizations in  $\sqrt{L}+\sqrt{C}$  (*-lec*) or  $\sqrt{L}+\sqrt{K(A)}$  (*-lka*) resist productively combining with the frequent verbalizer *-irati*, typically used in borrowed verbs (e.g. *analizirati* ‘analyse’, *promovirati* ‘promote’, etc.). Out of 125 verbs derived with *-irati* which figure among the 3000 most frequent verbs in the corpus Gigafida, only two have attested nominalizations of this type, and both of them are attested less than three times: *kopiralec* ‘copier’ and *parkiralec* ‘parker’. Second, there are nominalization in which  $\sqrt{L}+\sqrt{C}$  (or  $\sqrt{L}+\sqrt{K(A)}$ ) is preceded by verbal-looking material which is not attested in any actual verb. Such  $\sqrt{L}+\sqrt{C}$ -nominalizations are illustrated in (16).

- (16)  $\sqrt{L}+\sqrt{C}$ -nominalizations lacking the base verb
- | Nominalization                | Implied verb  | Actual verb                 | Nominalization of the actual verb |
|-------------------------------|---------------|-----------------------------|-----------------------------------|
| posnemovalec<br>‘imitator’    | *posnemovati  | posnemati<br>‘imitate’      | ?posnemalec<br>‘imitator’         |
| obračunovalec<br>‘calculator’ | *obračunovati | obračunavati<br>‘calculate’ | *obračunavalec                    |
| obveščevalec<br>‘informant’   | ?obveščevati  | obveščati<br>‘inform’       | *obveščalec                       |

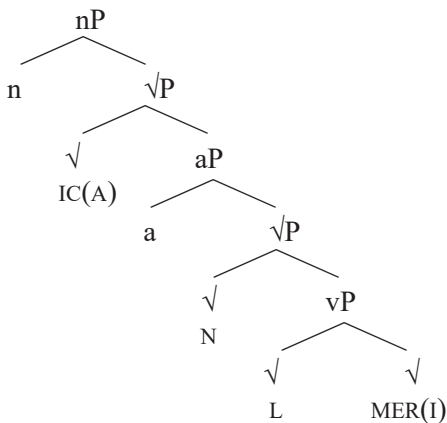
A remarkable feature of the nominalizations in (16) is that they all contain the sequence *ova/eva*, which is readily analysed as the root  $\sqrt{ov}$  in combination with the theme vowel *a*, a pattern widely attested in actual verbs (see Simonović/Mišmaš, this volume). So, similarly to the examples in (13b) and (13c),  $\sqrt{L}$  selects a structure which contains a theme vowel, but this structure is not necessarily an existing verb.

Before concluding this contribution, it should be pointed out that the model presented predicts the existence of words in which root-selecting and category-selecting affixes co-occur in the same structure. I have not presented any such words so far because the presentation has focused on the dichotomy between radical cores and their default stress, on the one hand, and structures which intertwined roots and functional structure with their faithful stress, on the other. However, words which have both radical cores and category-selecting affixes are amply attested in Slovenian. In (17) two such nouns are presented which are related to the verb *méri* ‘measure’.

- (17) Derivations containing *méri+l* related to *méri* ‘measure’ (part 2)
- |                  |                    |
|------------------|--------------------|
| merilnic-a       | merilnik           |
| ‘measuring room’ | ‘measuring device’ |

Note that these nominalizations do not have stem-final stress, but they crucially contain a radical core (*meriln*) which has received the default stress. The analysis is shown in (18).

(18)



## 5 CONCLUSION

In this contribution, a principled account has been presented of stress assignment in deadjectival and deverbal nominalizations in Slovenian, explicitly addressing both

regular cases and exceptions. The analysis has been couched in the model proposed by Lowenstamm (2014), which views derivational affixes as transitive roots. The main addition to the model was the proposal that parts of the structure which only contain roots with no intermediate functional structure (the ‘radical cores’) always receive default prosody. The analysis presented dispenses with lexical stress in Slovenian affixes and derives the prosodic properties of affixed words solely from their structure.

The proposed model has been applied only to a limited number of derivational patterns, but it makes rather clear predictions as to what is to be expected in the rest of the language(s). It is therefore a worthwhile endeavor for future research to test the proposed model using additional data both from Slovenian and other languages.

Finally, since the main focus has been on prosody, at least two aspects of the data analysed have received less attention than they deserve. The first is the semantics of the nominalizations discussed, which has only been addressed in passing. The second is the nature of the theme vowels in Slovenian (and beyond). For the sake convenience, I have followed Marvin (2003) in assuming that theme vowels belong to the root. It, however, remains a fact that theme vowels only play a role in analyses which relate mainly to the verbal domain. It is therefore necessary to explore further their nature and their lexical representation. It is my hope that further research will tackle these issues and bring new insights to bear on the matter.

## Primary Sources

Gigafida. 24 February 2020. <http://www.gigafida.net/>.

## References

- HALLE, Morris/Alec MARANTZ (1993) “Distributed morphology and the pieces of inflection.” In: K. Hale/S. J. Keyser (eds), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*. Cambridge, MA: MIT Press, 111–176.
- HALLE, Morris/Alec MARANTZ (1994) “Some key features of distributed morphology.” In: A. Carnie/H. Harley/T. Bures (eds), *MIT Working Papers in Linguistics* 21. Cambridge, MA: MITWPL, 275–288.
- LOWENSTAMM, Jean (2014) “Affixes as roots.” In: A. Alexiadou/H. Borer/F. Schäfer (eds), *The syntax of roots and the roots of syntax*. Oxford: Oxford University Press, 230–259.
- MARANTZ, Alec (1996) “‘Cat’ as a phrasal idiom.” Manuscript. Cambridge, MA: Massachusetts Institute of Technology.
- MARVIN, Tatjana (2003) *Topics in the stress and syntax of words*. Doctoral dissertation. Cambridge, MA: Massachusetts Institute of Technology.
- REVITHIADOU, Anthi (1999) *Headmost accent wins: Head dominance and idealprosodic form in lexical accent systems*. Doctoral dissertation. Leiden: Leiden University.
- SIMONOVIĆ, Marko (2020) “Derivational affixes as roots across categories.” [In press.]

Abstract  
CATEGORIES, ROOT COMPLEXES AND DEFAULT STRESS:  
SLOVENIAN NOMINALIZATIONS REVISITED

A new account is presented of stress assignment in deadjectival and deverbal nominalizations in Slovenian, explicitly addressing both regular cases and exceptions. The analysis is an extension of the account by Marvin (2003) and is couched in the model developed by Lowenstamm (2014), which views derivational affixes as transitive roots. The main addition to Lowenstamm's model is the proposal that parts of the structure which only contain roots with no intermediate functional structure (the 'radical cores') always receive default prosody. The presented analysis dispenses with lexical stress in Slovenian affixes and derives the prosodic properties of affixed words solely from their structure.

**Keywords:** stress, nominalizations, Distributed Morphology, default stress pattern, derivational affixes as roots, radical cores

Povzetek  
KATEGORIJE, KORENSKI SKUPKI IN PRIVZETI NAGLAS:  
PONOVO O SLOVENSKIH NOMINALIZACIJAH

V prispevku je predstavljen nov pristop k določevanju naglasnega mesta izpridevniških in izglagolskih nominalizacij, ki neposredno naslavlja tako običajne kot izjemne primere. Pristop nadgrajuje analizo Tatjane Marvin (2003) z modelom Jeana Lowenstamma (2014), v katerem so besedotvorni morfemi obravnavani kot prehodni morfemi. Glavni prispevek k Lowenstammovemu modelu je predlog, v skladu s katerim deli strukture, ki vsebujejo le korene brez vmesne funkcijske strukture (t. i. korenski skupki), vedno pridobijo privzeto prozodijo. Predstavljena analiza tako odpravlja leksikalni naglas v slovenskih morfemih in razlaga prozodične lastnosti tvorjenk izključno z vidika njihove strukture.

**Ključne besede:** naglas, nominalizacije, razpršena morfologija, privzeti naglasni vzorec, besedotvorni morfemi kot koreni, korenski skupki



**Susanne Wurmbrand\***  
University of Vienna, Austria

UDK 811.163'367.625:81'362  
DOI: 10.4312/linguistica.60.1.119-137



**Iva Kovač\*\***  
University of Vienna, Austria

**Magdalena Lohninger\*\*\***  
University of Vienna, Austria

**Caroline Pajančič\*\*\*\***  
University of Vienna, Austria

**Neda Todorović\*\*\*\*\***  
University of British Columbia, Canada

## FINITENESS IN SOUTH SLAVIC COMPLEMENT CLAUSES: EVIDENCE FOR AN IMPLICATIONAL FINITENESS UNIVERSAL

### 1 INTRODUCTION

The notion of finiteness is a much-debated topic in syntax, morphology, and semantics, as many fundamental questions have not been answered yet. Due to the vast cross-linguistic variation in the distribution of finiteness, many works have concluded that there is no single morpho-syntactic definition of finiteness, nor a single semantic function associated with it (see e.g., Cristofaro 2007; Bisang 2007; Nikolaeva 2007). Among the morpho-syntactic categories that have been suggested to reflect finiteness are tense, aspect, mood, illocutionary force, person marking, politeness, special forms not used in independent clauses, and/or nominal morphology on the verb (see the works in Nikolaeva 2007). Semantically, finiteness is often associated with clausal independence, specifically, the possibility of a sentence to occur as a free-standing (declarative) main clause (Maas 2004; Bisang 2007; Givón 1990), or, particularly in the *Government and Binding* and *Minimalism* traditions, with independent tense or an anchoring to the logophoric center of a clause (Bianchi 2003; Adger 2007). Despite the (theoretical and empirical) variation, various distributional tendencies and patterns have been observed (see Givón 1990; Cristofaro 2007), which indicate that the distribution is not arbitrary but follows certain implicational relations.

In this paper, we will look at variation in the distribution of finiteness in a well-defined empirical domain – the South Slavic languages (SSL), and show that it falls

---

\* susanne.wurmbrand@univie.ac.at

\*\* iva.kovac@univie.ac.at

\*\*\* magdalena.lohninger@univie.ac.at

\*\*\*\* caroline.pajancic@univie.ac.at

\*\*\*\*\* neda.todorovic@ubc.ca

along two dimensions: language and type of complement. To illustrate, while Serbian (Sr) allows finite and non-finite complements of verbs like *try* (1a), Slovenian (Sl) only allows non-finite complements in this context (1b).<sup>1</sup> On the other hand, when the matrix verb is a speech verb like *claim* in (1c), the distribution in Slovenian is exactly the opposite from (1b) – only a finite complement is possible.

(1) a. *Pokušala sam {da čitam / čitati} ovu knjigu.* (Sr)  
 tried.SG.F AUX.1SG {DA read.1SG / read.INF.IPFV} this book  
 ‘I tried to read this book.’

b. *Poskusila sem {\*da berem / brati} to knjigo.* (Sl)  
 tried.SG.F AUX.1SG {\*da read.1SG / read.INF.IPFV} this book  
 ‘I tried to read this book.’ (Adrian Stegovec, p.c.)

c. *Trdim, {da berem / \*brati} to knjigo.* (Sl)  
 claim.1SG {da read.1SG / \*read.INF.IPFV} this book  
 ‘I claim that I am reading/to be reading this book.’ (Adrian Stegovec, p.c.)

The property in (1c) holds for all SSL and we propose that the uniformity in the choice of finite forms in these types of complements is due to a grammatical constraint shared by all SSL. We also show that in addition to the two types of complements in (1), there is a third class, illustrated by complements to verbs like *decide* in (2), which shows flexibility regarding finiteness in most, but not all SSL languages.

(2) *Odločila sem se {brati /da bom brala} to knjigo.* (Sl)  
 decided.SG.F AUX.1SG REFL {read.INF.IPFV /da will.1SG read.SG.F}  
 this book  
 ‘I decided to read this book.’ (Adrian Stegovec, p.c.)

Comparing constructions like (1) and (2) across the SSL, we argue that complement clauses form a semantic hierarchy, and that the distribution of (non-)finiteness in the SSL reflects an implicational scale along this hierarchy. We suggest that the implicational nature of the hierarchy (and the distribution of finiteness) is derived via containment relations of clausal domains. Lastly, we propose that the variation is the result of different distributions of the features triggering finiteness, specifically in SSL, agreement features. Our findings thus support the existence of implicational hierarchies, defined in semantic terms, on which morphological coding, such as manifestations of finiteness, operates.

More broadly, we will conclude that: i) there is no (universal) semantic correlate of (non-)finiteness, since cross-linguistically, all types of complements can be realized as

1 Since we will argue that *da* does not always act as a complementizer, we gloss it as DA in the examples.



finite or non-finite in at least some language; ii) finiteness is not confined to a particular domain in the clause, but can be distributed over all clausal domains (see also Adger 2007); but iii) there are systematic implicational relations which hold among different types of complements and which, despite possibly arbitrary specific finiteness settings, allow us to predict certain properties of finiteness.

## 2 IMPLICATIONAL COMPLEMENTATION HIERARCHY (ICH)

Languages exhibit a variety of different types of complementation, which can be divided into different classes based on their semantic and/or their morphosyntactic properties, such as finite/non-finite, subjunctive, nominalization (see among others Givón 1980; Pesetsky 1992; Horie 2001; Cristofaro 2005; Dixon 2006; Noonan 2007 for different approaches). This paper follows Givón's basic insight that classes are defined semantically, forming an implicational complementation hierarchy (ICH), which we define below. Syntactic and morphological distinctions (such as finiteness), if present in a language, operate along that scale. While morphosyntactic effects may be neutralized in different languages, the semantic scale is observable cross-linguistically, and morphosyntactic distinctions can never go against the hierarchy. Wurmbrand and Lohninger (to appear) propose that in addition to the fine-grained semantic scale given in Givón (1980), languages bundle categories into three super-sets, which we refer to as *Proposition*, *Situation*, and *Event*, adopting the terminology and definitions in Ramchand and Svenonius (2014) (a similar classification has been proposed in Rochette 1988, 1990, although with different terminology). Complements of the type *Proposition* involve speech, epistemic, and factive contexts. These types of complements are temporally independent, have no pre-specified tense value, are anchored in an utterance or embedding context, and may involve speaker-oriented parameters. Complements of the type *Situation* involve emotive and irrealis contexts. These types of complements are elaborate eventualities without speaker- and utterance-oriented properties, but with time and world parameters, allowing them to refer to a specific, possibly pre-determined, time. The most common type of *Situation* complements are forward expanded unrealized events where the time of the complement must be after the time of the matrix verb (Abusch 2004; Wurmbrand 2014b). Complements of the type *Event* include implicative and strong attempt contexts. These types of complements are semantic *Properties* (Chierchia 1984; Wurmbrand 2002) in that they lack speaker- and utterance-oriented, as well as time and world parameters; they are tenseless, may involve actuality entailments (Hacquard 2006), and may have reduced argument structure and/or event properties. Using this classification, Wurmbrand (2014a, 2015), and Wurmbrand and Lohninger (to appear) establish the implicational nature of the ICH as in Table 1. *Independence* refers to properties such as the presence and/or interpretation of an independent subject in the complement clause; *transparency* characterizes whether the embedded clause is permeable for certain cross-clausal operations or dependencies; and *integration* refers to the degree to which the embedded predicate is incorporated into the matrix predicate.

Table 1: Implicational complementation hierarchy (ICH)

MOST INDEPENDENT LEAST TRANSPARENT LEAST INTEGRATED	Proposition $\gg$ Situation $\gg$ Event	LEAST INDEPENDENT MOST TRANSPARENT MOST INTEGRATED
---	---	--

ICH hierarchy effects have originally been discussed predominantly for infinitives, where clause union or restructuring effects are easily detectable. However, in Todorović and Wurmbrand (2020) and Wurmbrand and Lohninger (to appear), it is observed that even in languages with no infinitives, complements show (in)dependence and transparency effects that track the ICH. Moreover, the semantic classes identified above are typologically robust, whereas morphosyntactic properties differ significantly across languages. If, as we suggest, the basic complementation classes are defined semantically by the ICH, we expect to see the different (semantic) types of complements to display variable properties, even when morphosyntactic finiteness does not distinguish between them. That the ICH applies to finite and non-finite complements alike is illustrated here by the distribution of clause-introducers in complement clauses in Bulgarian and Macedonian, both languages that do not have infinitives. As shown in (3a), *Proposition* complements must occur with *če* in Bulgarian and cannot be introduced by *da*. *Situation* complements as in (3b) can occur with either *če* or *da*. Lastly, *Event* complements can only occur with *da*, as in (3c).

- (3) a. *Lea tvārdi*                    {*če* / \**da*}    *čete*                    *kniga*.  
 Lea claim.PRS.3SG {that / \**da*} read.PRS.3SG book  
 ‘Lea claims that she is reading a book.’                    (Marchela Oleinikova, p.c.)
- b. *Lea reši*                        {*če* \*(*šte*) / *da* }    *čete*                    *kniga*.  
 Lea decided.PRF.3SG {that \*(will) / *da* } read.PRS.3SG book  
 ‘Lea decided to read/that she will read a book.’
- c. *Lea se opitvaše*                {\**če* / *da*}    *čete*                    *kniga*.  
 Lea REFL try.PRF.3SG {\*that / *da*} read.PRS.3SG book  
 ‘Lea tried to read a book.’

The same restrictions hold for Macedonian as illustrated in (4).

- (4) a. *Lea tvārdi*                    {*deka* / \**da*}    *čita*                    *kniga*.  
 Lea claim.PRS.3SG {that / \**da*} read.PRS.3SG book  
 ‘Lea claims that she is reading a book.’                    (Sandra Jakimovska, p.c.)
- b. *Lea se rešila*                    {*deka* \*(*e*) / *da* }    *čita*                    *kniga*.  
 Lea REFL decided.PRF.3SG {that \*(will) / *da* } read.PRS.3SG book  
 ‘Lea decided to read/that she will read a book.’

- c. *Lea probala* {*\*deka* / *da*} *čita* *kniga*.  
 Lea try.PRF.3SG {*\*that* / *da*} read.PRS.3SG book  
 ‘Lea tried to read a book.’

As shown in (3b), while *da* complements can encode future directly, *če/deka* complements require an overt future element to be interpreted as a *Situation* complement (see section 4.1 for an explanation of this difference). In Bulgarian, some speakers may also allow a *decide* complement introduced by *če* and no overt future. In this case, the configuration typically cannot receive a *Situation* interpretation, but is instead shifted to a *Proposition* context such as a performative use where the matrix subject evaluates or assigns truth to the embedded proposition (similar to cases like *I decided that he is a nice person*; but other attitude meanings are possible for some speakers as well). Thus, predicates may sometimes undergo a class change, and it is therefore essential to pair the morphosyntactic properties with the interpretation when evaluating ICH effects. If a *Situation* predicate can be changed or coerced into a *Proposition* predicate, it follows from the ICH (which is defined semantically) that the predicate then takes on the morphosyntactic properties of the *Proposition* class (such as the complementizer *če*). The distribution of complement clauses in Bulgarian (Bu) and Macedonian (Ma) is summarized in Table 2.

Table 2: Clause introducers in Bu, Ma

	Proposition	Situation	Event
<i>če/deka</i>	✓	✓ (with future)	*
<i>da</i>	*	✓	✓

As shown in Table 2, Bulgarian and Macedonian exhibit what we refer to as an *ICH signature effect*: the two classes at the opposite ends of the ICH scale show the opposite properties, while the class in the middle is ‘in-between’ in that it is compatible with both properties (in this case). In the next section, we will show that the distribution of finiteness in SSL shows ICH signature effects and follows the ICH in an interesting way. In section 4, we then suggest a direction for deriving the ICH and the finiteness distribution. Section 5 concludes the paper.

### 3 FINITENESS HIERARCHY

#### 3.1 A Finiteness ICH Signature Effect

We now turn to the SSL that have infinitives, i.e., Bosnian, Croatian, Slovenian, and Serbian. To illustrate that SSL show ICH signature effects regarding (non-)finiteness, we begin by summarizing the finiteness property of the three types of complements in Slovenian, as shown in (5) (repeated from (1) and (2); see below for further explanations regarding (5b)).

- (5) a. *Trdim, {da berem / \*brati} to knjigo.* (Sl)  
 claim.1SG {da read.1SG / \*read.INF.IPFV} this book  
 ‘I claim that I am reading/to be reading this book.’
- b. *Odločila sem se {brati / da bom brala / (\*da berem} to knjigo.*  
 decided.SG.F AUX.1SG REFL {read.INF.IPFV / da will.1SG read.SG.F /  
 (\*) da read.1.sg} this book  
 ‘I decided to read this book.’
- c. *Poskusila sem {brati / \*da berem / \*da bom brala} to knjigo.*  
 tried.SG.F AUX.1SG {read.INF.IPFV / \*da read.1SG / \*da will.1SG  
 read.SG.F} this book  
 ‘I tried to read this book.’

As in all SSL, *Proposition* complements can only be finite (5a). *Situation* complements can be either non-finite or finite. In Slovenian (and many other languages), finite complements usually require an overt future tense to convey the same meaning as the infinitive in (5b), unless the language and context allows future interpretations to be expressed by present tense (e.g., as in planned or scheduled events). If a present-for-future use is not possible, for whatever reasons, a finite form without overt future (i.e., the third option in (5b)) is excluded when the intended reading is the same as the reading of the corresponding infinitival construction. Similarly to Bulgarian, however, this form may still be rendered acceptable if the matrix verb is shifted to a performative (i.e., a *Proposition*) interpretation (e.g., (5b) could be used in a situation where the subject is making a decision about which book to read in a play and then declares which book they decided on). Finally, Slovenian *Event* complements, as shown in (5c), can only occur as non-finite. These observations are summarized in Table 3.

Table 3: Finiteness in Slovenian complements

	<b>Proposition</b>	<b>Situation</b>	<b>Event</b>
<i>finite</i>	✓	✓ (with future)	*
<i>non-finite</i>	*	✓	✓

Comparing the distribution of Bulgarian and Macedonian clause introducers with the finiteness distribution in Slovenian, we see an obvious parallel. Both cases display a clear ICH signature effect: the *Proposition* and *Event* classes show opposing properties, whereas the *Situation* class allows both (albeit with certain restrictions).

### 3.2 A Possible Finiteness Universal

In the broader context of South Slavic, different languages show different patterns of availability of infinitives and finite forms; the distribution of these forms seems to be largely dependent on the geographical location and language contact. As can be seen in Figure 1, Bulgarian and Macedonian do not allow infinitives at all, Croatian is the most infinitive-friendly language, and Bosnian, Slovenian, and Serbian occupy the middle of the scale; Slovenian inclines more towards the non-finite, Serbian towards the finite extreme of the scale, and Bosnian is in the middle (judgements for Bosnian differ, thus illustrating the language contact situation quite well).

NO INFINITIVES	Bu/Ma >> Se >> Bo? >> Sl >> Cr	FREQUENT INFINITIVES
----------------	--------------------------------	----------------------

Figure 1: Infinitives in the SSL

Since Bulgarian and Macedonian do not use infinitives, we concentrate here on the grammatical patterns of the other four SSL, Bosnian (Bo), Croatian (Cr), Slovenian (Sl), and Serbian (Sr). The distribution is given in (6). Note that due to the contact situation of these languages, language/dialect affiliation is not always clear-cut. Since categorical judgments may not always be possible, the marks should be understood as preferences. The data are given in Serbian in (6) (see Vrzić 1996), but the judgments are to be understood as applying to the translations of these examples into the different languages.

- (6) a. *Tvrđim {da čitam / čitati} ovu knjigu.*  
 claim.1SG {da read.1SG / read.INF.IPFV} this book  
 ‘I claim that I am reading this book.’ (finite) √Sr, √Bo, √Sl, √Cr  
 ‘I claim to be reading this book.’ (non-finite) \*Sr, \*Bo, \*Sl, \*Cr
- b. *Odlučila sam {da čitam / da ću čitati / čitati} ovu knjigu.*  
 decided.SG.F AUX.1SG {da read.1SG / da will.1SG read / read.INF.IPFV}  
 this book  
 ‘I decided that I will read this book.’ (finite) √Sr, √Bo, √Sl, \*/?Cr  
 ‘I decided to read this book.’ (non-finite) ?Sr, √Bo, √Sl, √Cr
- c. *Pokušala sam {da čitam / čitati} ovu knjigu.*  
 tried.SG.F AUX.1SG {da read.1SG / read.INF.IPFV} this book  
 ‘I tried that I am reading/will read this book.’ (finite) √Sr, ?Bo, \*Sl, \*Cr  
 ‘I tried to read this book.’ (non-finite) √Sr, √Bo, √Sl, √Cr

As shown in (6a), the SSL are uniform with *Proposition* complements in that they disallow infinitives across the board. The *Situation* class in (6b) allows infinitives in

all SSL (even though they are dispreferred in many Serbian varieties), but otherwise exhibits variation. Serbian (and possibly Slovenian)<sup>2</sup> can express the future meaning with a finite present tense form. Bosnian, Slovenian, and Serbian allow infinitives or finite overt future forms, but Croatian strongly disprefers any finite form.<sup>3</sup> Lastly, the *Event* class in (6c) permits infinitives in all SSL, with it being the only possible form in Croatian and Slovenian. A finite complement clause is allowed in Serbian and possibly in Bosnian, but judgements in Bosnian differ and we have not been able to conclusively determine the distribution. The distribution of finite vs. non-finite complements is summarized in Table 4 (due to the variation within Bosnian, we have not been able to conclusively allocate it to a category, and we therefore list it in two places). The distribution clearly shows that there is a variation according to the two parameters: language and type of complement.

Table 4: Finiteness in South Slavic

	<b>Proposition</b>	<b>Situation</b>	<b>Event</b>
<i>Bulgarian, Macedonian</i>	finite	finite	finite
<i>Serbian, Bosnian?</i>	finite	(non-)finite	(non-)finite
<i>Slovenian, Bosnian?</i>	finite	(non-)finite	non-finite
<i>Croatian</i>	finite	non-finite	non-finite

The tendencies observed in the SSL reveal a clear finiteness scale which follows the ICH: a type of complement can never be ‘more’ finite than the type of complement to its left. On the basis of this distribution we propose the (hypothetical) finiteness universal in (7).

(7) (Hypothetical) Finiteness Universal

If a language {allows/requires} finiteness in a type of complement, all types of complements further to the left on ICH also {allow/require} finiteness.

## 4 A SKETCH OF AN ACCOUNT

### 4.1 Towards Deriving the ICH

While the ICH in Givón 1980 is defined functionally, Wurmbrand and Lohninger (to appear) propose a grammatical approach. To address the question of how the ordering and implicational nature of the ICH arise, we start with the mapping of the semantic sorts *Proposition*, *Situation*, and *Event* to syntax. Following Ramchand and Svenonius (2014), *Propositions*, *Situations*, and *Events* are semantic sorts expressing conceptual

2 An anonymous reviewer suggested that this may be possible in Slovenian as well, but we have not investigated yet whether this is restricted to present-for-future contexts or possible for future statements in general.

3 This has been confirmed by speaker judgements, preliminary corpus searches on Google, and a query on hrWaC (the Croatian web corpus).

primitives which are in a coherent containment relation – *Situations* are elaborations of *Events*, *Propositions* are elaborations of *Situations*. More specifically, *Situations* are created by combining time/world parameters with an existentially closed *Event*, and *Propositions* combine speaker-oriented/discourse-linking parameters with an existentially closed *Situation*. In other words, these semantic sorts are computed in a predictable way by combining the verb with its arguments (creating an *Event*), by relating an (existentially closed) *Event* to a time through T or other temporal elements (creating a *Situation*), and by anchoring a *Situation* to a context through an element of the operator domain (e.g., C, creating a *Proposition*). Note that while the broad distinction into clausal domains is considered a general property of phrase structure (see also Grohmann 2003), the detailed internal organization of these domains (e.g., CP, TMA, and vP) may vary cross-linguistically. Figure 2 illustrates the containment relations among clausal domains and their semantic correspondences.

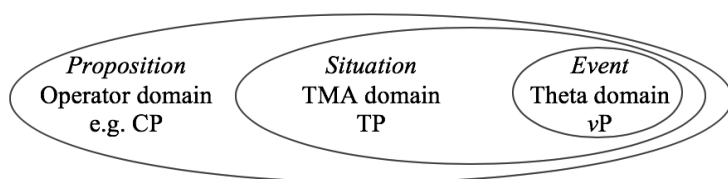


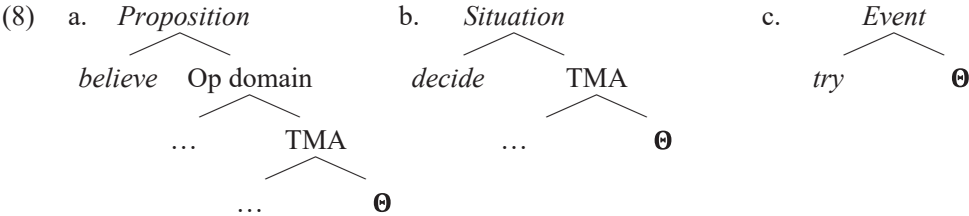
Figure 2: Clausal domains

Since, as we propose, complementation is also defined by the semantic sorts *Proposition*, *Situation*, and *Event*, the same containment implications arise. As specified in Table 5, the three types of complements have different *minimal requirements*. *Proposition* complements are cognition and utterance complements with independent and not predetermined tense interpretations (*I said that he left/will leave/is leaving*). There is still a tense dependency in the sense that the tense in every complement clause is interpreted relative to the event time of the matrix predicate, but, crucially, the choice of the embedded tense value is free. Following Kratzer (2006) and Moulton (2009a,b), aspects of the meaning of an attitude configuration are situated in the operator domain of the complement clause. The operator domain also separates the matrix predicate and the embedded temporal domain, leading to the tense value independence of these complements. *Situation* complements involve an independent temporal domain in that the matrix and embedded temporal interpretation can differ (e.g., *I decided yesterday to leave tomorrow*). However, the tense value is predetermined – the embedded clause, whether infinitive or finite, should be situated after the event time of the matrix predicate. There is thus a closer/tighter connection between the matrix predicate and the embedded temporal domain. Lastly, *Event* complements do not involve an independent tense – they are always interpreted as simultaneous with the matrix predicate. The ranking and implicational nature of the ICH can then be seen as a reflex of the resulting semantic complexity scale.

Table 5: Clausal domains and complexity

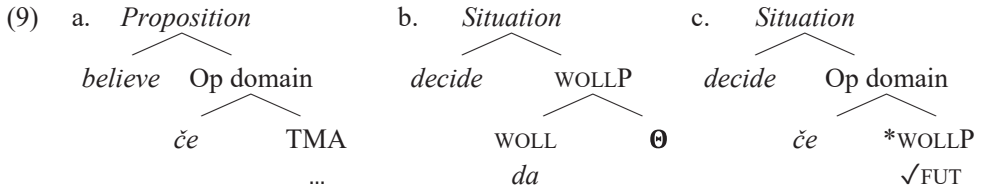
	Proposition	Situation	Event
Minimal requirement	Operator domain TMA domain Theta domain	TMA domain Theta domain	Theta domain
Complexity	most complex	intermediate	least complex

In this paper, we will pursue the hypothesis that the different types of complements can vary in minimal size as in (8). This does not mean that *Event* complements are necessarily always just theta domains (e.g., *vPs*). Structures could vary across languages, however, our main claim is that they can never go against the hierarchy. For instance, there could not be a language where *Proposition* complements are always less complex than *Event* or *Situation* complements.



Returning to the distribution of clause introducers in Bulgarian, the system outlined here can derive the ICH signature effect as well as the restriction noted about overt future. The assumption we make is that *če* is a true complementizer (i.e., a head in the operator domain), whereas *da* is a lower clausal head. Since *Proposition* complements require the operator domain, it follows that they always occur with *če*, as in (9a). *Situation* complements can occur without the operator domain, in which case the embedded clause, a TMA domain, directly combines with the matrix verb. The irrealis interpretation arises through a covert future modal *WOLL* (Wurmbrand 2014b; Todorović 2015), which, following Todorović and Wurmbrand (2020), needs to be identified within the syntactic context. One way to license *WOLL* is via Merge with a *Situation* verb as in (9b) (we leave open whether this is a selectional or featural relation). If, on the other hand, a *Situation* verb combines with an operator domain, (9c), the matrix verb and *WOLL* are too far apart, and only an independent overt future is possible (in which case *WOLL* is licensed by Tense (Abusch 1988)). The resulting future statement, although structurally more complex than a simple *WOLL*-clause, still satisfies the semantic requirement of a *Situation* verb, which only demands that the complement refers to a forward expanded unrealized event.





In the next section, we show how this implementation of the ICH derives the finiteness universal in (7) (repeated in (10)).

(10) (Hypothetical) Finiteness Universal

If a language {allows/requires} finiteness in a type of complement, all types of complements further to the left on ICH also {allow/require} finiteness.

**4.2 Finiteness in the SSL**

An account of the distribution of finiteness in the SSL in Table 4, repeated as Table 6, needs to derive the following properties:

- i) *Proposition* complements are always finite in the SSL (but not cross-linguistically).
- ii) *Situation* and *Event* complements can be finite or non-finite in the SSL, with language-specific settings restricting the options.
- iii) The distribution of finiteness follows the implicational universal in (7).

Table 6: Finiteness in South Slavic

	<b>Proposition</b>	<b>Situation</b>	<b>Event</b>
<i>Bulgarian, Macedonian</i>	finite	finite	finite
<i>Serbian, Bosnian?</i>	finite	(non-)finite	(non-)finite
<i>Slovenian, Bosnian?</i>	finite	(non-)finite	non-finite
<i>Croatian</i>	finite	non-finite	non-finite

Before answering these questions, we need to address one basic, yet very difficult question, namely what finiteness is.

**4.2.1 What is Finiteness?**

Typically, finiteness is associated with a property of the temporal domain (but see Cristofaro 2007; Bisang 2007 for other options). Comparing the distribution of finiteness in Table 6 with the temporal properties of these complements in Table 7, it becomes clear that there is no general semantic tense property that can be mapped to a morphosyntactic category finite or non-finite in SSL.

Table 7: Embedded tense properties

	<b>Proposition</b>	<b>Situation</b>	<b>Event</b>
<i>Embedded temporal domain</i>	yes	yes	no
<i>Predetermined tense value</i>	no	yes	N/A

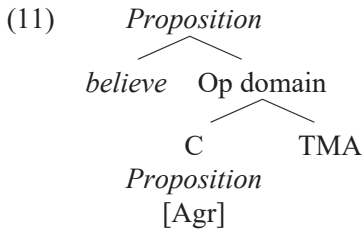
While it may be tempting to treat the uniformity of finiteness in *Proposition* complements as a reflex of a semantic property (e.g., independent, non-predetermined tense; see also below), this would raise the question of why *Situation* complements can (Serbian, Bosnian, Slovenian) or must (Bulgarian, Macedonian) also be finite, and why even *Event* complements can (Serbian, Bosnian) or must (Bulgarian, Macedonian) be finite. It would also be difficult to extend such an account to languages outside the SSL, where *Proposition* complements can also be non-finite (e.g., English *She claims to have won*). Similarly, although *Event* complements tend to be non-finite, this cannot be the result of mapping the semantic lack of tense to a morphosyntactic category non-finite, since it is only a trend – Serbian and Bosnian allow, and Bulgarian, Macedonian require finite *Event* complements, where, importantly, the interpretation is the same as in *Event* contexts in the other languages. Lastly, *Situation* complements also show the entire spectrum from obligatory finite (Bulgarian, Macedonian), optionally finite (Serbian, Bosnian, Slovenian), to non-finite (Croatian), where, once again, the interpretation of the complement in all languages is the same, namely that of a forward expanded unrealized event.

While we conclude that it is not possible to directly derive the distribution of finiteness from semantic properties, we have seen in section 4.1 that the dependencies in Table 7 nevertheless play an important role in the distribution of complementation itself. The different temporal properties reflect a scale of independence of the embedded clauses, which we have suggested is structurally implemented via different syntactic complexities. This approach allows us to derive the implicational nature of the ICH, and it will also provide an answer for the question of why the distribution of finiteness follows the implicational universal in (7).

Returning to the question of what finiteness is, we follow Cristofaro (2007) and Bisang (2007), who argue that there is no universal category of finiteness, but that languages can differ in what properties they utilize to express finiteness. We suggest that in the SSL, finiteness corresponds to agreement (whereas in other languages it could be tense, the combination thereof, or even other properties). More specifically, we follow Adger (2007), where it is suggested that features related to finiteness are not confined to a particular syntactic position (such as a Fin head in the CP), but can also occur on lower clausal heads. For instance, Adger suggests that subject licensing in Scottish Gaelic (a finiteness property) requires the (uninterpretable) features [T] and [Agr], but these features do not necessitate a semantic T or C head – they can also be inserted (somewhat parasitically) on other heads or project independent AgrPs in any domain of the clause. Although the details of finiteness in Scottish Gaelic and the SSL differ, we adopt this proposal in spirit and assume that in the SSL, finiteness is the spell-out of agreement features, which can occur on  $\nu$ , T, or C.

#### 4.2.2 Proposition Complements

We are now able to derive the distribution of finiteness in complement clauses in SSL, starting with *Proposition* complements. We have observed that, whenever the operator domain must be projected, i.e., in *Proposition* complements (see Table 5), finiteness is obligatory (and this holds regardless of the finiteness settings of *Event* and *Situation* complements). We propose that in *Proposition* complements, the locus of finiteness is the CP in SSL. More specifically, as illustrated in (11), if the operator domain is projected in a complement clause, agreement features are obligatory and they occur on C or a split CP with a separate agreement projection (we do not split Agr in the trees).



Finiteness in the CP is, of course, what one would expect following cartographic approaches. Our approach differs, however, in several respects. First, as we will see momentarily, finiteness is not restricted to the CP domain. In particular, we argue, following Adger (2007), that finiteness cannot entail the presence of a CP. Adger proposes that finiteness comes in two versions – a semantic notion of finiteness, associated with the CP, and occurrences of finiteness in the lower clausal domains which are not associated with a semantic function of finiteness. While we adopt the second part of Adger’s proposal, we submit that finiteness in complement clauses is never associated with a semantic property. Following Bianchi (2003), Adger suggests that a semantic finiteness head in the CP is responsible for identifying the embedded event/reference time [E/R] with the speech time [S] and/or the relation of participants to the external logophoric center, the external speech event. While this is a possible approach for main clauses, it does not carry over to embedded clauses. Bianchi (2003: 7) already qualifies the claim that “A finite verb form can encode the relation of E/R to S” with “at least in main clauses.” In complement clauses, the embedded tense is not related to the speech time, but always evaluated with respect to the matrix tense. Thus, even in *Proposition* complements, which have their own temporal domain and no predetermined tense value, there is a tense dependency with the matrix clause and not the speech event. Furthermore, according to Bianchi (2003: 7), “A non-finite form does not encode any relation to S”. Since this is true for all types of complement clauses (none of them involve deictic tense), it would not allow us to distinguish between the different types of complements, and it therefore offers no way to approach the distribution in Table 6. The only option to tie finiteness in *Proposition* complements in the SSL to a semantic property is to associate it with the attitude property itself (following Kratzer 2006;

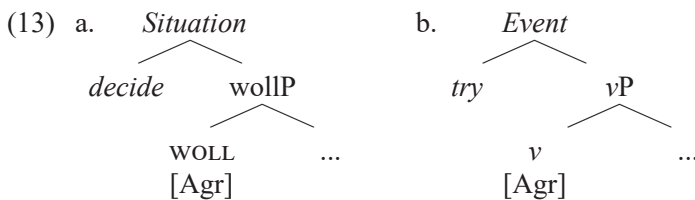
Moulton 2009a,b). But since *Proposition* complements are not finite universally (see e.g., English, German), it once again cannot be a necessary connection between a semantic and morphosyntactic property. Our approach captures this – finiteness is not semantic but purely morphosyntactic. However, the distribution of this morphosyntactic property, implemented by the projection of agreement features, is sensitive to the structure, which in turn is determined by semantic properties. For the SSL we thus have the language-specific property in (12).

(12) *C* [*Proposition*]: +Agr

The last point to note is how an Agr head/feature in the CP triggers finiteness on the verb. There are various technical ways to derive this. For the sake of simplicity, we assume that the clausal heads *C*, *T*, *v+V* enter a dependency with each other (this could be V/T-movement, Agree, or feature lowering) and that the highest verbal element must realize the Agr feature.

#### 4.2.3 *Situation* and *Event* Complements

Since *Situation* and *Event* complements can lack the operator domain (see Progovac 1993a,b, 1994, 1996; Stjepanović 2004; Todorović and Wurmbrand 2020 for detailed motivation for clause reduction despite finiteness), under a cartographic approach where finiteness is located (solely or partially) in the CP, the question arising is how clauses without CP can be finite. Our account can successfully capture the presence of finiteness in the absence of CP as well as the distribution in Table 6. Following Adger (2007), we assume that agreement features, which are responsible for finiteness in the SSL, can be located in the other clausal domains, i.e., on heads of the TMA domain and even on heads of the theta domain. The structure of a finite *Situation* complement is given in (13a), the one of a finite *Event* complement in (13b).



The distribution of agreement features in the TMA and theta domains must be restricted by language-specific settings, which are spelled out in Table 8 for the SSL.

Table 8: Agr distribution

	CP domain	TMA domain	Theta domain
<i>Bulgarian, Macedonian</i>	Agr	Agr	Agr
<i>Serbian, Bosnian?</i>	Agr	optional Agr	optional Agr
<i>Slovenian, Bosnian?</i>	Agr	optional Agr	no Agr
<i>Croatian</i>	Agr	no/marginal Agr	no Agr

The approach thus captures the fact that *Situation* and *Event* complements can be finite or non-finite, even though *Proposition* complements must always be finite in the SSL. Since the finiteness of *Proposition* complements comes from the CP, and the CP can be missing in non-*Proposition* complements, it is correctly predicted that finiteness is only obligatory (in the SSL) when the operator domain is required. In all other types of complements, non-finite forms could be possible due to different settings of the lower clausal heads, which may or may not come with an Agr head. The approach thus allows us to derive the cross-linguistic variation – languages could differ in the inventory of Agr associated with heads of the different domains. The last advantage of this approach is that, despite the language-specific settings in Table 8, there are predictions it makes, which we turn to in the last subsection.

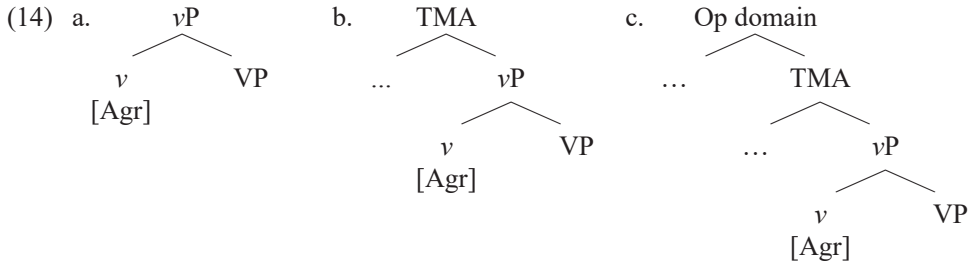
#### 4.2.4 Finiteness Universal

Recall that the distribution of finiteness in different types of complements is systematic and follows the finiteness universal – if a complement is allowed/required to be finite in a language, all the complements to the left of it allow/require finiteness. Thus, for instance, there is no language where *Event* complements are obligatorily finite and *Situation* and/or *Proposition* complements optionally non-finite; or no language where *Event* complements are optionally finite but *Situation* and/or *Proposition* obligatorily non-finite. Note that this does not mean that it is not possible for a language to realize finite *Event* complements and non-finite *Situation* complements at the same time – but what is impossible is that these be the only options in the language.

Our approach derives these implicational relations in the distribution of finiteness. Since clausal domains are in a containment configuration (see Figure 2), it follows that settings in a lower domain affect all clauses that include that domain, i.e. also clauses with additional higher domains, since higher domains necessarily include the lower ones. To be more concrete, if in a language the theta domain is specified as (obligatorily) finite (Bulgarian, Macedonian), all types of clauses will be realized as finite since the theta domain is included in all clause types.<sup>4</sup> In other words, if *Event* complements can or must be finite in a language, all other types of complements in the language can

<sup>4</sup> The specification for Agr in Table 8 would thus not be necessary for the TMA and CP domains in these languages, since the Agr feature in the theta domain is sufficient to trigger finiteness in all types of clauses. We included it in the table for expository purposes, but also to leave open the option of double finiteness marking (see Todorović and Wurmbrand 2020).

or must be finite as well. This is illustrated in (14) – finiteness as in (14a) entails finiteness in configurations (14b,c). Although the Agr specifications for heads can differ in our system, it is not possible to derive a configuration in which, for instance, the theta domain contains Agr but clauses that include higher domains do not. Even if one were to posit Agr on *v*, but no Agr specifications on TMA or CP heads, the containment relations predict that lower Agr specifications can never be ‘undone’.



## 5 CONCLUSION

We have shown in this paper that the distribution of (non-)finiteness in the SSL reflects an implicational scale along an independently attested semantic hierarchy. We have suggested that in the SSL, finiteness is triggered by clausal agreement features associated with different syntactic heads. Building on a complexity approach to the complementation hierarchy, cross-linguistic variation in finiteness, as well as variation across different types of complements is derived as the result of language-specific differences in the distribution of agreement features. More broadly, we have concluded that there is no (universal) semantic correlate of (non-)finiteness and, contra cartographic approaches, that finiteness is not confined to a particular domain in the clause, but can be distributed over all clausal domains.

## Acknowledgements

We thank the audience of the Workshop on the *Morphology of South Slavic languages* and an anonymous reviewer for invaluable feedback. This work has been supported by the Austrian Science Fund (FWF): M 2332-G30 (Universals and variation in clausal complementation) and P34012-G (Implicational hierarchies in clausal complementation).

## References

- ABUSCH, Dorit (1988) “Sequence of Tense, intensionality and scope.” In: H. Borer (ed), *The Proceedings of the 7<sup>th</sup> West Coast Conference on Formal Linguistics*. Stanford: CSLI Publications, 1–14.
- ABUSCH, Dorit (2004) “On the temporal composition of infinitives.” In: J. Guéron/J. Lecarme (eds), *The syntax of time*. Cambridge, MA: Massachusetts Institute of Technology, 27–53.

- ADGER, David (2007) “Three domains of finiteness: A minimalist perspective.” In: I. Nikolaeva (ed), 23–58.
- BIANCHI, Valentina (2003) “On finiteness as logophoric anchoring.” In: J. Guéron/L. Tasmovski (eds), *Temps et point de vue/Tense and Point of View*. Paris: Université Paris X. 213–246.
- BISANG, Walter (2007) “Categories that make finiteness: Discreteness from a functional perspective and some of its repercussions.” In: I. Nikolaeva (ed), 115–137.
- CHIERCHIA, Gennaro (1984) *Topics in the syntax and semantics of infinitives and gerunds*. Doctoral dissertation. Amherst: University of Massachusetts.
- CRISTOFARO, Sonia (2005) *Subordination*. Oxford: Oxford University Press.
- CRISTOFARO, Sonia (2007) “Deconstructing finiteness: Finiteness in a functional-typological perspective.” In: I. Nikolaeva (ed), 91–114.
- DIXON, Robert M. W (2006) “Complement clauses and complementation strategies in typological perspective.” In: R. M. W Dixon/A. Y. Aikhenvald (eds), *Complementation: A cross-linguistic typology*. Oxford: Oxford University Press, 1–48.
- GIVÓN, Talmy (1980) “The binding hierarchy and the typology of complements.” *Studies in Language* 4/3, 333–377. <https://doi.org/10.1075/sl.4.3.03giv>
- GIVÓN, Talmy (1990) *Syntax: A Functional-Typological Introduction, Vol. II*. Amsterdam: J. Benjamins.
- GROHMANN, Kleanthes K. (2003) *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam/Philadelphia: John Benjamins.
- HACQUARD, Valentine (2006) *Aspects of modality*. Doctoral dissertation. Cambridge, MA: Massachusetts Institute of Technology.
- HORIE, Kaoru (2001) “Complement clauses.” *Language Typology and Language Universals. An International Handbook* 2, 979–993.
- KRATZER, Angelika (2006) “Decomposing Proposition verbs.” *Talk given in honor of Anita Mittwoch*. The Hebrew University of Jerusalem.
- MAAS, Utz (2004) “‘Finite’ and ‘nonfinite’ from a typological perspective.” *Linguistics* 42/2, 359–385. <https://doi.org/10.1515/ling.2004.012>
- MOULTON, Keir (2009a) “Clausal complementation and the *wager*-class.” In: A. Schardl/M. Walkow/M. Abdurrahman (eds), *Proceedings of the North Eastern Linguistics Society Annual Meeting 38*. Amherst: University of Massachusetts, GLSA, 165–178.
- MOULTON, Keir (2009b) *Natural selection and the syntax of clausal complementation*. Doctoral dissertation. Amherst: University of Massachusetts.
- NIKOLAEVA, Irina (2007a) “Constructional economy and nonfinite independent clauses.” In: I. Nikolaeva (ed), 138–180.
- NIKOLAEVA, Irina (ed) (2007b) *Finiteness: Theoretical and empirical foundations*. Oxford: Oxford University Press.
- NIKOLAEVA, Irina (2007c) “Introduction.” In: I. Nikolaeva (ed), 1–19.
- NOONAN, Michael (2007) “Complementation.” In: T. Shopen (ed), *Language typology and syntactic description: Complex Constructions (Volume II)*. Cambridge: Cambridge University Press, 52–150.

- PESETSKY, David (1992) *Zero syntax II: An essay on infinitives*. Cambridge, MA: MIT.
- PROGOVAC, Ljiljana (1993a) “Locality and subjunctive-like complements in Serbo-Croatian.” *Journal of Slavic Linguistics* 1, 116–144.
- PROGOVAC, Ljiljana (1993b) “Subjunctive: The (mis)behavior of anaphora and negative polarity.” *The Linguistic Review* 10, 37–59. <https://doi.org/10.1515/tlir.1993.10.1.37>
- PROGOVAC, Ljiljana (1994) *Negative and positive polarity: A binding approach*. Cambridge: Cambridge University Press.
- PROGOVAC, Ljiljana (1996) “Clitics in Serbian/Croatian: Comp as the second position.” In: A. Halpern/A. Zwicky (eds), *Approaching second: Second position clitics and related phenomena*. Stanford: CSLI Publications, 441–428.
- RAMCHAND, Gillian/Peter SVENONIUS (2014) “Deriving the functional hierarchy.” *Language sciences* 46, 152–174. <https://doi.org/10.1016/j.langsci.2014.06.013>
- ROCHETTE, Anne (1988) *Semantic and syntactic aspects of Romance sentential complementation*. Doctoral dissertation. Cambridge, MA: Massachusetts Institute of Technology.
- ROCHETTE, Anne (1990) “On the restructuring classes of verbs in Romance.” In: A. Di Sciullo/A. Rochette (eds), *Binding in Romance: Essays in honour of Judith McA’Nulty*. Ottawa: Canadian Linguistic Association, 96–128.
- STJEPANOVIĆ, Sandra (2004) “Clitic climbing and restructuring with ‘finite clause’ and infinitive complements.” *Journal of Slavic Linguistics* 12/1–2, 173–212.
- TODOROVIĆ, Neda (2015) “Tense and aspect (in)compatibility in Serbian matrix and subordinate clauses.” *Lingua* 167, 82–111. <https://doi.org/10.1016/j.lingua.2015.04.007>
- TODOROVIĆ, Neda/Susi WURMBRAND (2020) “Finiteness across domains.” In: T. Radeva-Bork/P. Kosta (eds), *Current Developments in Slavic Linguistics. Twenty Years After (based on selected papers from FDSL 11)*. Frankfurt am Main: Peter Lang, 47–66.
- VRZIĆ, Zvezdana (1996) “Categorial status of the Serbo-Croatian ‘modal’ da.” In: J. Toman (ed), *Formal Approaches to Slavic Linguistics: The College Park Meeting 1994*. Ann Arbor: Michigan Slavic Publications, 291–312.
- WURMBRAND, Susi (2002) “Syntactic versus semantic control.” In: C. J. Zwart/W. Abraham (eds), *Studies in comparative Germanic syntax*. Amsterdam: John Benjamins, 95–129.
- WURMBRAND, Susi (2014a) “Restructuring across the world.” In: L. Veselovská/M. Janebová (eds), *Complex Visibles Out There. Proceedings of the Olomouc Linguistics Colloquium 2014: Language Use and Linguistic Structure*. Olomouc: Palacký University, 275–294.
- WURMBRAND, Susi (2014b) “Tense and aspect in English infinitives.” *Linguistic Inquiry* 45/3, 403–447. [https://doi.org/10.1162/LING\\_a\\_00161](https://doi.org/10.1162/LING_a_00161)
- WURMBRAND, Susi (2015) “Restructuring cross-linguistically.” In: T. Bui/D. Özyıldız (eds), *Proceedings of the North Eastern Linguistics Society Annual Meeting 45*. Amherst: University of Massachusetts, GLSA, 227–240.



WURMBRAND, Susi/Magdalena LOHNINGER (to appear) “An implicational universal in complementation – Theoretical insights and empirical progress.” In: J. M. Hartmann/A. Wöllstein (eds), *Propositional Arguments in Cross-Linguistic Research: Theoretical and Empirical Issues*. Berlin: Mouton de Gruyter.

#### Abstract

#### FINITENESS IN SOUTH SLAVIC COMPLEMENT CLAUSES: EVIDENCE FOR AN IMPLICATIONAL FINITENESS UNIVERSAL

This paper shows that the distribution of (non-)finiteness in the South Slavic languages reflects an implicational scale along an independently attested semantic complementation hierarchy (e.g., Givón 1980). We suggest that in the South Slavic languages, finiteness is triggered by clausal agreement features associated with different syntactic heads. Building on a complexity approach to the complementation hierarchy, we propose that cross-linguistic variation in finiteness and variation across different types of complements are the result of language-specific differences in the distribution of agreement features. More broadly, we conclude that there is no (universal) semantic correlate of (non-)finiteness and, contra cartographic approaches, that finiteness is not confined to a particular domain in the clause. Following Adger (2007), we argue that finiteness can be distributed over all clausal domains.

**Keywords:** syntax, morphology, language variation, implicational hierarchies, complementation, infinitives, finiteness, South Slavic

#### Povzetek

#### OSEBNE IN NEOSEBNE GLAGOLSKE OBLIKE V JUŽNOSLOVANSKIH DOPOLNILIH: DOKAZI ZA IMPLIKACIJSKO UNIVERZALIJO

V članku pokažemo, da porazdelitev osebnih in neosebnih glagolskih oblik v južnoslovanskih jezikih odraža implikacijsko lestvico po neodvisno potrjeni semantični hierarhiji dopolnil (npr. Givón 1980). Zagovarjamo trditev, da v južnoslovanskih jezikih pojavnost (ne)osebnih glagolskih oblik sprožajo stavčne ujemalne oznake, povezane z različnimi skladenjskimi jedri. Na podlagi pristopa kompleksnosti k hierarhiji dopolnil predlagamo, da so medjezikovne razlike v pojavnosti (ne)osebnih glagolskih oblik in razlike med vrstami dopolnil rezultat jezikovno specifičnih razlik v porazdelitvi ujemalnih oznak. V splošnem zaključujemo, da ne obstaja (univerzalni) pomenski korelat pojavnosti (ne)osebnih glagolskih oblik in da, v nasprotju s kartografskim pristopom, ta lastnost ni omejena na specifično domeno v stavku. Podobno kot Adger (2007) trdimo, da se lastnost (ne)osebnih glagolskih oblik lahko porazdeli po vseh stavčnih domenah.

**Ključne besede:** skladnja, morfologija, jezikovna variacija, implikacijske hierarhije, dopolnila, nedoločniki, (ne)osebne glagolske oblike, južnoslovanski jeziki

LINGUISTICA LX  
(2020)

Založila  
Znanstvena založba Filozofske fakultete Univerze v Ljubljani  
Izdal  
Oddelek za romanske jezike in književnosti

Revue éditée par les  
Presses scientifiques de la Faculté des Lettres  
et publiée par le  
Département des Langues et Littératures Romanes

Za založbo – Responsable  
Roman Kuhar  
Dekan Filozofske fakultete – Doyen de la Faculté des Lettres

Glavna in odgovorna urednica – Rédactrice en chef  
Martina Ožbot

Številko LX uredili – Numéro LX dirigés par  
Christina Manouilidou, Tatjana Marvin, Gašper Ilc, Andrej Stopar

Lektoriranje – Relecture  
Oliver Currie

Tajnica redakcije – Secrétaire de rédaction  
Metka Šorli

Dopise nasloviti na:  
Prière d'adresser toute correspondance à :  
Martina Ožbot  
Filozofska fakulteta  
Oddelek za romanske jezike in književnosti  
Aškerčeva 2  
1000 Ljubljana  
Slovénie

[linguistica@ff.uni-lj.si](mailto:linguistica@ff.uni-lj.si)

Tel.: + 386 1 241 13 98

Fax: + 386 1 425 93 37

Naklada: 400 izvodov – Tirage : 400 exemplaires

Računalniški prelom – Mise en page  
Aleš Cimprič

Tisk – Impression  
Birografika BORI, d. o. o.  
Linhartova cesta 1, 1000 Ljubljana

Cena: 17 €