

Processing of Croatian Aspectual Derivatives

Krešimir Šojat¹, Kristina Kocijan²[0000-0001-9467-5313], Matea Filko¹

¹ Department of Linguistics

² Department of Information and Communication Sciences

Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb, Croatia
{ksojat, krkocijan, matea.filko}@ffzg.hr

Abstract. The main objective of this paper is to detect and describe major derivational processes and affixes used in the derivation of aspectually connected Croatian verbs. This kind of analysis is enabled by previous detection of verbal derivational families, i.e. families of verbs with the same root as well as the derivational affixes they contain. Using NooJ, we automatically detect such derivational processes and assign the aspectual tag to derivatives. The procedure is based on the list of selected base forms and derivatives, on the list of derivational affixes and their allomorphs, and on the set of derivational rules. For this objective we selected 15 verbal derivational families comprising app. 250 derivatives in total. The output is being used for the development of a large on-line database of Croatian aspectual pairs, triples and quadruplets. Such a resource will be valuable for various research works in lexicology and lexicography.

Keywords: Derivationally connected verbs, prefixation, suffixation, aspectual derivatives, aspectual pairs, aspectual triples, aspectual quadruplets, Croatian, NooJ.

1 Introduction

This paper deals with computational processing of Croatian derivational morphology. We focus on verbal derivation and aspect. Our objective is to present preliminary work done during the construction of the database of Croatian aspectually and derivationally connected verbs, i.e. **aspectual derivatives**.

Croatian is a South Slavic language with very rich inflectional and derivational morphology. Inflectional phenomena are extensively covered by two publicly available large lexica for Croatian – Croatian Morphological Lexicon (CML) (Tadić & Fulgosi, 2003) and hrLex (Ljubešić et al., 2016). Each lexicon, used for various NLP tasks such as lemmatization, POS and MSD tagging, etc., contains complete inflectional data for more than 100 000 lemmas. The computational processing of Croatian derivation is on a much smaller scale compared to the size of these inflectional lexica.

CroDeriV (Šojat et al., 2013) is an on-line database that contains app. 14 500 Croatian verbs and provides information about their morphological structure and derivational relatedness. Derivational families consist of verbs that share the same lexical

morpheme. Although CroDeriV enables the detection of all derivational affixes in lemmas, derivational processes, e.g. prefixation or suffixation, within derivational families are currently neither specified nor indicated to users. Still, CroDeriV is a valuable source of data for various research, including this one as well.

Here, we use linguistic data from CroDeriV for the detection of derivational processes and affixes within selected derivational families. By using NooJ (Silberstein, 2016) as our NLP tool, we firstly aim to automatically detect processes such as prefixation or suffixation. Secondly, we want to automatically assign the aspectual tag to derivatives, i.e. to determine whether a verb is perfective or imperfective.

The paper is structured as follows: in Section 2 we briefly describe major derivational processes in Croatian and focus on the derivation of verbs from other verbs and aspectual changes that take place. Section 3 deals with analysis of data, whereas in Section 4 the NooJ dictionary of verbs is presented. In section 5 we dissect the morphology to find the patterns that we can use for the NooJ grammar and provide an overview of underlying principles. In Section 6 the design and the structure of the web-based database of Croatian aspectual verbal pairs is briefly discussed. The paper concludes with an outline of the future work.

2 Derivational Processes and Aspectual Changes

Derivation and compounding are major word-formation processes used in Croatian. However, unlike in some other languages, e.g. German, compounding is not as productive as derivation. For the purposes of this project, we will deal only with derivation, which is in Croatian mainly based on affixation. Our main focus is the derivation of verbs from other verbs. Although there are some other processes, like conversion and back formation, they are not as prominent as for example prefixation and suffixation.

As far as derivationally connected verbs are concerned, i.e. those that share the lexical morpheme and therefore belong to the same derivational family, they are derived from other verbs via prefixation, suffixation, stem alternations or various combinations of these processes. We only briefly deal with stem alternations (cf. Section 3), since this area requires a different approach due to frequent allomorphy of roots or stems. Still, a complete list of variants is a prerequisite for an accurate description.

Full derivational spans of selected base forms in terms of verb-to-verb derivation used in this paper are extracted from CroDeriV. The derivational span refers to all derivatives that are connected to a particular base form. The base form refers to the simplest verb within a family regarding its morphological structure and is used for the derivation of other members. The size of verbal derivational families significantly varies: some of them consist of only one or two members, while others encompass more than 30 or 40 derivatives. For example, the derivational family based around the base form *pisati* ‘to write’ contains 31 verbal derivatives. Out of this number, more than 50% (16 in total) of verbs are derived via prefixation.

As in other Slavic languages, each verb in Croatian is always marked for aspect and classified as perfective, imperfective, or bi-aspectual. Generally, the perfective

aspect is used to describe actions, processes and states as finished or completed, whereas the imperfective aspect refers to them as unfinished or ongoing, e.g.:

1. a. *Pisala je* [imperfective] članak jedan sat. 1b. She *was writing* an article for an hour.
2. a. *Napisala je* [perfective] članak za jedan sat. 2b. She *wrote* an article in an hour.

Verbs like *pisati* ‘to write + imperfective’ – *napisati* ‘to write, to finish writing + perfective’ are usually referred to as aspectual pairs. Verbs in aspectual pairs are closely related in meaning, except that one expresses perfective and the other imperfective aspect. Aspect in Croatian is inherent verbal category – it is morphologically marked in each verbal form and it affects inflectional properties of verbs to a certain degree (Kocijan et al., 2018). Furthermore, it is regarded as a word-formation process and members of aspectual pairs are treated as separate lexical entries in dictionaries.

Although the verbal aspect in Slavic languages is based on the opposition of only two aspects and it is overtly marked, numerous studies in the area of second language acquisition indicate that aspect is one of the most complicated category for learners of Slavic languages (Cvikić & Jelaska, 2007).

In terms of derivation, perfectives are commonly derived from imperfectives by prefixation, while imperfectives can be formed from perfectives by suffixation or stem alternation. The presence of certain affixes indicates whether a verb is a perfective or an imperfective. A relatively small group of bi-aspectual verbs, mostly of foreign origin, can be used as perfectives and imperfectives in the same morphological form. Various factors can determine whether they will be used as perfectives or imperfectives (e.g. a context, the type of time adverbial used in a sentence, etc.).

As indicated, prefixation is the most productive process in the derivation of verbs from other verbs, although other affixes enable further derivation, either through multiple prefixation, suffixation or simultaneous prefixation and suffixation. Croatian verbs can thus be divided into simple imperfectives (*pisati* ‘to write + imperfective’) for on-going actions and prefixed perfectives (*na-pisati* ‘to write + perfective’) for completed actions. Such pairs are referred to as **primary aspectual pairs**. Further derivation of perfectives in primary aspectual pairs is not possible.

It is important to notice that other prefixes used for the derivation of perfectives in this derivational family can add different semantic features to the meaning of the base verb (e.g. *pisati* ‘to write + imperfective’ – *pre-pisati* ‘to copy by writing + perfective’ – *pot-pisati* ‘to sign + perfective’) as thoroughly discussed in Šojat et al. (2012). In such cases, further derivation of aspectual derivatives is possible, either through prefixation, suffixation or simultaneous prefixation and suffixation. Polančec (2018) explains how such perfectives can be derived into secondary imperfectives usually denoting iterative actions through suffixation (*potpis-iva-ti* ‘to sign several/many times’).

Simultaneous prefixation and suffixation yields derivatives usually denoting actions performed in a sufficient, abundant or excessive manner like in the following examples:

- *jesti* ‘to eat + imperfective’ – *najesti se* ‘to eat one’s fill + perfective’;
- *raditi* ‘to work + imperfective’ – *naraditi se* ‘to tire oneself out with work + perfective’;

- *pisati* ‘to write + imperfective’ – *napisati se* ‘to be weary of writing + perfective’.¹

On the other hand, a set of suffixes is used for the derivation of diminutive verbs as is the case with

- *pisati* – *pis-kar-a-ti* ‘to scribble + imperfective’

or verbs expressing punctual actions

- *vikati* ‘to shout + imperfective’ – *vik-nu-ti* ‘to shout once + perfective’.

Some secondary imperfectives are further derived via prefixation into perfectives denoting distributive actions as in

- *is-potpisivati* ‘to sign each one + perfective’, e.g. each letter, every document, etc.

On top of that, aspectual distinctions are in some cases expressed by vowel variations or suppletive forms, like

- *doći* ‘to come + perfective’ – *dolaziti* ‘to come + imperfective’.

To sum up, verbs in Croatian are derived from other verbs by prefixation and suffixation. Both processes can trigger a change in aspect and the addition of a new semantic component to the base form. Apart from aspectual change, semantic components brought by affixes can produce combinations that, in terms of meaning, can vary from compositional to completely idiosyncratic. Detailed account of such morpho-semantic relations among Croatian verbal derivatives, frequently referred to also as Aktionsart, is found in Šojat et al. (2012).

In the following section, the analysis of selected derivational families is presented. This analysis should enable the automatic detection of derivational processes and aspectual changes as well as the development of morphological rules used by NooJ in the later stages of this research (cf. Section 5).

3 Data Analysis and Rules

In order to learn how we can automatically detect and annotate major derivational processes and affixes used in the derivation of aspectually connected verbs, a thorough analysis was performed. This analysis was facilitated by the extraction of derivational families, i.e. families of verbs with the same root, as well as the derivational affixes they contain. We started the analysis by selecting 15 verbal derivational families from CroDeriV comprising app. 250 derivatives in total. We manually analyzed all derivational processes in these families, marked derivational chains and inserted aspectual tags to derivatives.

¹ Although the element *se* is normally regarded as a reflexive particle and not treated as an affix in Croatian textbooks on this subject, for the sake of demonstration we treat it here as a suffix.

To demonstrate, we shall use the derivational family grouped around the base form *pisati* ‘to write + imperfective’. This family group consists of 32 members.² The analysis consisted of assigning tags for aspect (IPF – imperfective, PF – perfective, BI – bi-aspectual) and specifying the type of affixation (prefixation, suffixation, etc.):

- *pisati* (IPF) – *napisati* (PF) – [prefixation]
- *pisati* (IPF) – *dopisati* (PF) – [prefixation]
- *dopisati* (PF) – *dopisivati* (IPF) – [suffixation]
- *dopisivati* (IPF) – *dopisivati se* (IPF) – [suffixation + se]
- *pisati* (IPF) – *potpisati* (PF) – [prefixation]
- *potpisati* (PF) – *potpisivati* (IPF) – [suffixation]
- *potpisivati* (IPF) – *ispotpisivati* (PF) – [prefixation]
- *pisati* (IPF) – *napisati se* (PF) – [prefixation + se]

We manually marked 250 derivatives from 15 derivational families following the same procedure. On the basis of this analysis, we developed 10 general rules that can be used for automatic detection of particular derivational processes and possible change of aspect. The rules are designed for pairs of verbs, while the description of full derivational paths (e.g. *pisati* – *prepisati* – *prepisivati* – *isprepisivati*) is in the testing phase. Based on the analyzed data, we have formed the following set of ten IF-THEN rules:

1. **IF** base form is prefixed
THEN simple imperfective →³ prefixed perfective (*pisati* – *dopisati*)
2. **IF** prefixed perfective is suffixed
THEN prefixed perfective → suffixed secondary imperfective (*dopisati* – *dopisivati*)
3. **IF** base form is suffixed
THEN simple imperfective → suffixed (deminutive/pejorative) imperfective (*pisati* – *piskarati*)
4. **IF** base form is suffixed with *-nu-*
THEN simple imperfective → suffixed (punctual) perfective (*vikati* – *viknuti*)
5. **IF** base form is simultaneously prefixed and suffixed
THEN simple imperfective → prefixed/suffixed imperfective (*pisati* – *napisati se*)
6. **IF** suffixed secondary imperfective is suffixed with *se*
THEN suffixed secondary imperfective → suffixed secondary imperfective + *se* (*dopisivati* – *dopisivati se* ‘to correspond in writing’)
7. **IF** prefixed perfective is prefixed
THEN prefixed perfective → multiple prefixed perfective (*dopisati* – *nadopisati*)
8. **IF** multiple prefixed perfective is suffixed
THEN multiple prefixed perfective → suffixed secondary imperfective (*nadopisati* – *nadopisivati*)

² Complete data for this derivational family and other discussed in the paper can be retrieved from <http://croderiv.ffzg.hr/>.

³ → stands for ‘changes into’.

9. **IF** multiple prefixed secondary imperfective is prefixed

THEN multiple prefixed secondary imperfective → multiple prefixed (distributive) perfective (*potpisivati* – *ispootpisivati*).

10. **IF** simple perfective is prefixed

THEN simple perfective → prefixed perfective (*baciti* – *izbaciti*).

For the derivation of aspectual pairs based on suffix alternation, e.g.:

- *bacati* ‘to throw + imperfective’ – *baciti* ‘to throw + perfective’ or
- *lupati* ‘to hit + imperfective’ – *lupiti* ‘to hit + perfective’.

an additional rule was introduced. Moreover, as mentioned in Section 2, these rules do not cover root or stem alternations used in verbal derivation. Such instances require a separate set of rules based on a list of detected allomorphs, as is the case in the following sets:

- *dovoditi* ‘to bring + imperfective’ – *dovesti* ‘to bring + perfective’.
- *gađati* ‘to aim + imperfective’ – *pogoditi* ‘to hit + perfective’.

Due to frequently unpredictable derivational paths, the rules can also not tackle derivations of verbs like:

- *čekati* ‘to wait + perfective’ – *očekivati* ‘to await + imperfective’ or
- *raditi* ‘to work + perfective’ – *suradivati* ‘to cooperate + imperfective’.

The problem is a missing link in a derivation from a simple imperfective to a prefixed secondary imperfective (prefixed perfectives like **suraditi* or **očekati* do not exist (Polančec, 2018)). Thus, such examples have to be tagged manually.

In Section 5 we further discuss how the morphological structure of verbs affects aspectual tagging of lemmas and how this problem can be handled with NooJ. Before that, we briefly present how aspectual data was incorporated into NooJ dictionary of Croatian verbs and how it benefited in terms of its enlargement and enrichment.

4 Dictionary of Croatian Verbs

NooJ language resources for Croatian (Vučković, 2009; Vučković et al., 2010; Kocijan et al., 2018) include a dictionary of verbs that holds 4 225 entries⁴. All the verbs have been marked for part of speech (V), paradigm responsible for generation of verbs’ flective forms and recently for aspect (Kocijan et al., 2018) as well. Morphological grammar uses the aspect information in order to properly annotate derived forms. Thus, it was important to add information on aspect directly to the dictionary entries.

⁴ The dictionary is continuously updated with new verbs. Thus, the number of main entries may vary from any previous and future references to this dictionary.

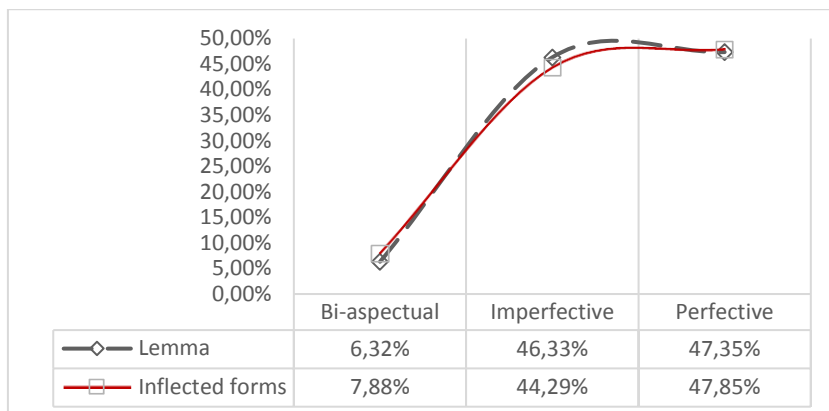


Fig. 1. Distribution of verbs by aspect in main (lemma) and fleective dictionaries (inflected forms)

Aspect wise, there are 267 bi-aspectual verbs marked as [+Aspect=dual], 1 957 imperfective verbs [+Aspect=inf], and 2 000 perfective verbs [+Aspect=fin]. When linked to its paradigm rules, NooJ produces dictionary of 376 583 fleective verb forms that holds 29 675 bi-aspectual, 166 753 imperfective and 180 155 fleective entries. Distributions of Aspects in both dictionaries (main and fleective) are almost identical (Fig. 1). The slight difference is due to the variation in number of tenses used in paradigm descriptions (Kocijan et al., 2018).

5 Dissecting Morphology

Croatian grammars (Babić 2002, Barić et al. 2003, Silić & Pranjković 2005) provided us with the list of prefixes and suffixes used for derivation of verbs. However, that was not enough for our project. We needed to make additional understanding of what happens before and after the main verb, and in some cases, inside the word. For this purpose, as stated earlier, a list of 15 base verbs with all their verbal derivatives was prepared. The selected base verbs are:

<i>pisati</i> ‘to write’	<i>raditi</i> ‘to work’	<i>bacati</i> ‘to throw’	<i>hraniti</i> ‘to feed’
<i>jesti</i> ‘to eat’	<i>piti</i> ‘to drink’	<i>kuhati</i> ‘to cook’	<i>čistiti</i> ‘to clean’
<i>čekati</i> ‘to wait’	<i>plakati</i> ‘to cry’	<i>ljubiti</i> ‘to kiss’	<i>reći</i> ‘to say’
<i>trčati</i> ‘to run’	<i>puzati</i> ‘to crawl’	<i>plivati</i> ‘to swim’.	

The selection of verbs was made mostly arbitrary trying to cover as much diversity as possible. The list was transferred to a sandbox area for thorough analysis that resulted with 8 distinguishable patterns that we will refer to as models 1 through 8.

The main difference between models is the number and position of affixes used for the derivation (Table 2). Prefixes are marked with letter P and a number 1 through 4 describing its position from the main root. Suffixes are marked with letter S and a number 1 through 3. In both cases, the larger the number, the farther away from the

root affixes are found. Each model has at least one prefix (P1) and one suffix (S3). If there is a prefix in position four (P4), all the proceeding positions must be filled as well. The opposite is true for suffixes, i.e. if there is a suffix in position one (S1), all the following positions must be filled (in this case S2 and S3).

Table 1. Eight derivational models

<i>Model</i>	Prefixes				Root	Suffixes		
	P4	P3	P2	P1	<i>root</i>	S1	S2	S3
M1	*	*	*	*	*		*	*
M2		*	*	*	*		*	*
M3		*	*	*	*			*
M4			*	*	*	*	*	*
M5			*	*	*		*	*
M6			*	*	*			*
M7				*	*		*	*
M8				*	*			*

Although the list of prefixes is a closed set, their selection depends on their position (P1, P2, P3 or P4). So far, we have detected only two paths for prefixes that fill all four positions⁵ (Fig. 3. – Area 4). Each position is marked as global variable @P, @P2, @P3 and @P4 respectful of their position.

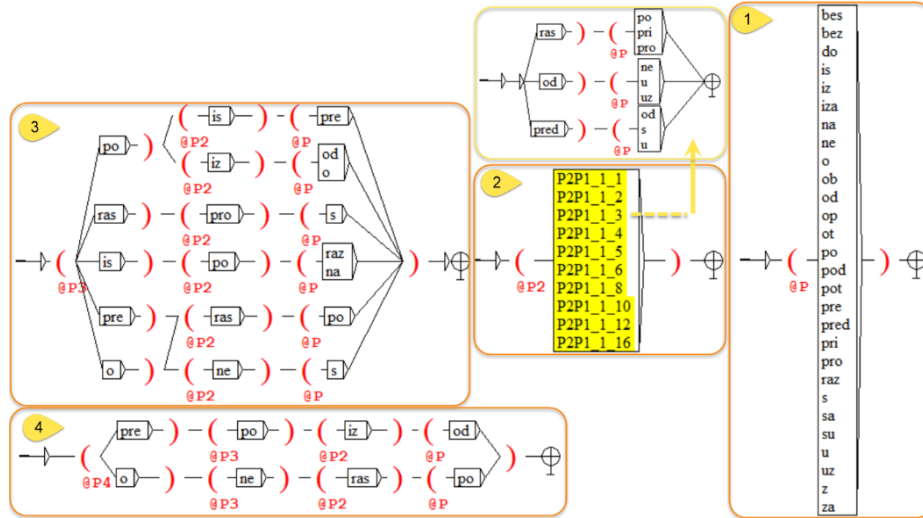


Fig. 2. Selection of prefixes depending on their number and position

⁵ In the present stage, we have treated the prefixes and their allomorphes as separate units in order to make the grammars in Nooj more simple and easier to process.

Prefixes that fill three positions (Fig. 3. – Area 3) have 9 possible combinations (paths), and depending on the context, some prefixes may be found in any of three positions (e.g.: *po-*), but they are never doubled or tripled. The same is true for prefixes that fill two positions (Fig. 3. – Area 2). However, the number of their combinations (paths) is much higher. In order to keep the grammars as clean and readable as possible, these combinations have been categorized into ten sets, depending on the number of P prefixes that can be found after the prefix in P2 position. For example, there are 14 P2 prefixes that can have only 1 (not necessarily the same) P prefix, but there are only 2 P2 prefixes (*po-*, *pre-*) that can have 16 P prefixes.

The matrix in Table 3 is used to show detected pairs when only two prefix positions are filled. Most of the prefixes are found in both positions, while five are found only in the first position i.e. in position P (*bes-*, *bez-*, *op-*, *sa-*, *z-*)⁶ and nine only in the second position i.e. in position P2 (*i-*, *poda-*, *pra-*, *pret-*, *raza-*, *re-*, *us-*, *ras-*).⁷ This matrix served us as a reference point for constructing the paths for each of the models.

Table 2. Ten categories of pairs of prefixes found in positions P2 (rows) and P1 (columns)

P2 P1 ->	bes	bez	do	is	iz	iza	na	ne	o	ob	od	op	ot	po	pod	pot	pre	pred	pri	pro	raz	s	sa	su	u	uz	z	za
i																												
iz																												
iza																												
ot																												
poda																												
pot																												
pra																												
pret																												
ra																												
raza																												
re																												
su																												
us																												
uz																												
do																												
ne																												
od																												
pred																												
ras																												
ob																												
pri																												
raz																												
su																												
na																												
pod																												
s																												
is																												
pro																												
iz																												
o																												
za																												
u																												
po																												
pre																												

⁶ Note that *bes-* and *bez-* are actually allomorphs of the prefix *bez-*, and that *sa-* and *z-* are allomorphs of the prefix *s-*. The prefix *i-* is actually an allomorph of the prefix *iz-*. Therefore, it would be correct to say that only three prefixes are found only in position P.

⁷ In this case, the prefix *i-* is an allomorph of the prefix *iz-*, *raza-* and *ras-* are allomorphs of the prefix *raz-*, *pret-* is an allomorph of *pred-*, *poda-* is an allomorph of *pod-* and *us-* is an allomorph of *uz-*.

After the learning phase during which we worked on the grammar design, we tested it on 1 650 verbs to see how well it performs. The selected verbs were either base verbs or derivatives, and are presently in our NooJ dictionary. The system scored 87% on both precision and recall. As expected, we found a number of false positives. Our future work will include their thorough analyzes and categorization, so we can learn from them in order to enhance the existing grammar.

6 The Database of Aspectual Derivatives

Behind the main initiative for this project is the development of an on-line database of Croatian aspectual pairs. We used NooJ to detect derivational processes and automatically assign aspectual tags to derivatives in a database suitable format. The procedure was based on two separate lists and a set of morphological rules described earlier. The first list is a selection of base forms and derivatives and the second one of derivational affixes and their allomorphs. The morphological rules that were designed have to perform two tasks: recognize the derived form and the verb it is derived from, and annotate the derived verb appropriately. All the tagged verbs can subsequently be automatically imported to the web-based database and used in a web search as described in Kocijan et al. (2018).

In its present form, the database provides the information about the main verb (including its aspect), aspect of a derived verb and affixes used within derivational family. In future, we plan to upgrade the database entries taking into consideration lexical semantics of base forms and derivatives. As it is indicated in Section 2, apart from aspectual change, derivations produce combinations that, in terms of meaning, vary from compositional to completely idiosyncratic. For example:

a. compositional:

- *trčati* ‘to run + imperfective’ – *utrčati* ‘to run into + perfective’ – *utrčavati* ‘to run into + secondary imperfective’
- *plivati* ‘to swim + imperfective’ – *uplivati* ‘to swim into + perfective’ – *uplivavati* ‘to swim into + secondary imperfective’

b. (more or less) idiosyncratic:

- *zreti* ‘to ripen + imperfective’ – *prezreti* ‘to scorn + perfective’ – *prezirati* ‘to scorn + secondary imperfective’
- *staviti* ‘to put + perfective’ – *predstaviti* ‘to introduce, to present + perfective’ – *predstavljati* ‘to introduce, to present + secondary imperfective’.

Our aim is to group prefixed perfectives and secondary imperfectives, as in examples above, into single entries within derivational families. The same principle will be followed for the aspectual combinations covered by the morphological rules in Section 3 of this paper.

However, due to complex semantics of Slavic verbs, sometimes seemingly simple and basic matters can turn out to be quite complicated. In many cases it is difficult to

choose or to decide even on a primary aspectual pair. For example, in the derivational family grouped around *raditi* ‘to work + imperfective’ there are two candidates (*uraditi* ‘to make + perfective’, *poraditi* ‘to make an effort + perfective’). Since these derivatives cannot be further suffixed, their choice would be fully justified in terms of morphological rules. However, semantically, none of the candidates corresponds to the base form, the meaning of derivatives is not compositional and, consequently, the entries for these aspectual derivatives cannot be generated automatically. We plan to further experiment with this line of work in future. Since a large amount of manual work is expected, the design of the database will enable a collaboration of multiple editors.

7 Conclusion and Future Work

We have presented preliminary stages in the construction of the database of Croatian aspectually and derivationally connected verbs, i.e. aspectual derivatives. Detected derivational models and categories of prefix’ combinations have been thoroughly described, visualized and exemplified to demonstrate the complexity of the project.

The morphological grammar proposed here lays down the fourfold basis for the following projects: a) detection of unknown verbs in the text and connecting them to the main (root) verbs they were derived from; b) automatic annotation of unknown verbs that will enlarge the existing NooJ dictionary of verbs (this should not be considered as the primary way of adding new verbs, but rather an auxiliary one and primary one being the regular dictionary input); c) usage of derivational strings as a learning tool (either for learners of Croatian as the primary or secondary language); d) export of annotated verbs to the web-based database. Such a database of Croatian aspectual derivatives is, to our knowledge, one of the first attempts to systematically present this area of Croatian derivational morphology. We believe that it will be a valuable resource for research not only in lexicology and lexicography, but in the domain of the second language acquisition, just the same.

References

1. Babić, S. Tvorba riječi u hrvatskome književnome jeziku. HAZU, Nakladni zavod Globus, Zagreb (2002).
2. Barić, E., Lončarić, M., Malić, D., Pavešić, S., Peti, M., Zečević, V., Znika, M. Hrvatska gramatika. Školska knjiga, Zagreb (2003).
3. Cvikić, L., Jelaska, Z.: Složenost ovladavanja glagolskim vidom u inojezičnome hrvatskome. Lahor 4, 190–216 (2007).
4. Kocijan, K., Šojat, K., Poljak, D.: Designing a Croatian Aspectual Derivatives Dictionary: Preliminary Stages. In: Proceedings of the First Workshop on Linguistic Resources for Natural Language Processing, pp. 28-37, Association for Computational Linguistics, Santa Fe, New Mexico, USA (2018).
5. Ljubešić, N., Klubička, F., Agić, Ž., Jazbec, I. P. New inflectional lexicons and training corpora for improved morphosyntactic annotation of Croatian and Serbian. In: Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016). ELRA, Portorož 4264–4270 (2016).

6. Polančec, J. Osamostaljeni izvedeni nesvršeni glagoli u hrvatskom jeziku. *Suvremena lingvistika*, 44 (85), 113–138.
7. Silberstein, M.: *Formalizing Natural Languages: The NooJ Approach*, Cognitive science series, Wiley-ISTE, London, UK (2016).
8. Silić, J., Pranjković, I. Gramatika hrvatskoga jezika za gimnazije i visoka učilišta. Školska knjiga, Zagreb (2005).
9. Šojat, K., Srebačić, M., Tadić, M. Derivational and Semantic Relations of Croatian Verbs. *Journal of Language Modelling*, 00 (1), 111–142 (2012).
10. Šojat, K., Srebačić, M., Štefanec, V. CroDeriV i morfološka raščlamba hrvatskoga glagola. *Suvremena lingvistika*, 39 (75), 75–96 (2013).
11. Tadić, M., Fulgosi, S. Building the Croatian Morphological Lexicon. In: *Proceedings of the EACL2003 Workshop on Morphological Processing of Slavic Languages* (Budapest 2003), ACL, 41–46 (2003).
12. Vučković, K.: *Model parsera za hrvatski jezik*. PhD dissertation, Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb (2009).
13. Vučković, K., Tadić, M., Bekavac, B.: Croatian Language Resources for NooJ. *CIT: Journal of computing and information technology*, 18(2010):295–301 (2010).