# Derivational and Semantic Relations of Croatian Verbs

*Krešimir Šojat*[1], *Matea Srebačić*[2], *and Marko Tadić*[1]
[1] University of Zagreb, Faculty of Humanities and Social Sciences, Zagreb
[2] University of Zagreb, Zagreb

## ABSTRACT

This paper deals with certain morphosemantic relations between Croatian verbs and discusses their inclusion in Croatian WordNet. The morphosemantic relations in question are the semantic relations between unprefixed infinitives and their prefixed derivatives. We introduce the criteria for the division of aspectual pairs and further discuss verb prefixation which results in combinations of prefixes and base forms that can vary in terms of meaning from compositional to completely idiosyncratic. The focus is on the regularities in semantic modifications of base forms modified by one prefix. The aim of this procedure is to establish a set of morphosemantic relations based on regular or re-occuring meaning alternations.

## 1    INTRODUCTION

In this paper we deal with certain types of derivational and semantic relations between Croatian verbs and discusss the possibilities of their inclusion in Croatian WordNet (CroWN), a lexical-semantic net built through the so-called expand model (cf. Vossen, 1998), i.e. by translating and adapting synsets from Princeton WordNet (PWN) into Croatian. At the present time CroWN consists of 10,000 synsets. Approximately 8500 of these are among the basic concept sets of EuroWordNet (EWN) and BalkaNet (BN). Most of the wordnets developed for other languages within these multilingual projects are based on the same or a very similar structure. Each synset in CroWN

was manually translated, adapted and provided with a definition of its meaning and several contextual usage examples. During this time-consuming work it became obvious that differences between Croatian and English are more significant than was initially assumed, especially when dealing with verbs. Despite the fact that PWN is used as a language-independent conceptual structure and as a sort of *interlingua*, some concepts are either not lexicalized in Croatian or are expressed in different parts of speech and therefore do not fit into lexical hierarchies as structured in PWN. For example, the English verb *to face*, as in *The two sofas face each other*, cannot be translated with a single verb in Croatian, and a construction like *to be opposite* should be used instead. Although multi-word literals are used in CroWN, this example and similar ones mainly relating to so-called stative verbs are problematic, since synsets contain words of the same part of speech. Multi-word units are units consisting of two or more words, but the whole unit should be the same POS as the other words in the same synset. However, the construction *to be opposite* consists of a verb and an adjective, and therefore, it is not a multi-word unit belonging to the lexical category of verbs. There are also numerous cases when verbs from PWN have both causative and reflexive translation equivalents in Croatian for the same synset (e.g., the verb *to melt* defined in PWN in one of its senses as *to become or cause to become soft or liquid*). Since all verbs in Croatian are always marked for aspect, the majority of English verbs have two or more translation equivalents in Croatian. The English verbs *drink* and *imbibe* — defined in PWN in one of their senses as *to take in liquids* and contextually illustrated with the sentence *The patient must drink several liters each day* — can be translated with at least five Croatian verbs: *piti, popiti, ispiti, ispijati, poispijati*[1]. Each of these verbs is morphologically derived from the basic verb *piti* 'to drink' through affixation, and each affix differently affects the base form in terms of lexical semantics. Whereas *popiti* denotes only that the action is finished, *ispiti* denotes that the action is finished and there is no liquid left. The lexical meaning of the verb *ispijati* includes both of these components as well as the additional com-

---

[1] Common strategies in dealing with aspect in Slavic wordnets are described in Section 2.

ponent denoting that the action is performed iteratively. Finally, the verb *poispijati* denotes that all these components are present, but the action is performed either by several subjects or on several objects.

The questions we pose here are (1) how to account for the derivational relations that exist between verbs in Croatian as an inflected language with rich derivation and (2) how to describe these relations in comparison to those which already exists among verbs in CroWN.

## 2                RELATED WORK

There are six main semantic relations between verbal synsets in PWN: synonymy, antonymy, proper inclusion, troponymy, presupposition and cause (cf. Fellbaum, 1998). As in EWN and BN, troponymy is substituted for hyponymy in CroWN, and cause is extended to encompass presupposition (cf. Vossen, 1998). Whereas the semantic relations in PWN connect synsets consisting of words from the same part of speech, in more recent work Fellbaum et al. (2007) discuss cross-POS relations that hold among words belonging to different synsets sharing a stem with the same meaning. These 14 "morphosemantic links", introduced into PWN 2.0., encompass relations based on suffixation patterns between verbs and nouns. Each relation is semantically labeled (e.g., Agentive, Instrument, Vehicle, Location etc.). None of the morphosemantic links include verb-verb pairs.

Pala and Hlaváčková (2007), Koeva (2008), and Koeva et al. (2008) discuss the problems they faced in the building of Czech, Bulgarian, and Serbian wordnets, respectively. Pala and Hlaváčková (2007) discuss the enrichment of Czech WordNet through the automatic generation of "derivational nests", i.e., new word forms derived from stems by adding affixes associated with specific meanings. They list 14 main derivational processes in Czech between nouns, verbs, adjectives, and adverbs, such as agentive relation in verb-noun pairs or diminutive relation in noun-noun pairs. Relations between derived and base form are semantically labeled and included in Czech Word-Net, resulting in a "two-level network". The higher level includes semantic relations between synsets such as synonymy, antonymy, or hyponymy. The lower level includes derivational relations between literals, i.e., single synset members. Verb-verb pairs are linked through

prefixation, but this relation is not taken into consideration in further processing and analysis.[2] Derivational relations are further classified into those which are predominantly semantic in nature, such as *agent*, and those which are predominantly morphological, such as *gerund*.

Koeva (2008) and Koeva et al. (2008) distinguish between morphosemantic and derivational relations. They claim that semantically related synsets in a source language for which the connection has been established through derivation can be used to link the synsets in a target language where such derivational links do not exist in the building of wordnets for several languages inter-connected via e.g., Interlingual index. The former, morphosemantic relations, are not language-specific, whereas the latter, derivational mechanisms of lexicalization, are language-specific. The sharing of semantic information across wordnets on similar grounds has also been proposed by Bilgin (2004) in the building of Turkish WordNet. Koeva (2008) stresses that one of the most productive derivational relations in Bulgarian is between verbal aspectual pairs and points out that perfective and imperfective verbs in Bulgarian WordNet are to be split into separate synsets that are subordinate to the same immediate hypernym. The relation of hypernymy would be based on imperfective verbs only. Derivationally related aspectual verb pairs would therefore be linked as literals. On the other hand, synsets would be linked with the morphosemantic relation *aspect*. The work presented in Koeva et al. (2008) concerning derivational and morphosemantic relations in Serbian WordNet is based on the same grounds and refers mainly to derivational relations across different parts of speech. In Serbian WordNet, aspectual pairs are members of the same synset. Each literal is provided with additional information about its inflectional and aspectual properties. The authors point out that, apart from aspectual pairs, perfective verbs derived from imperfective verbs by prefixation often have a different meaning and thus are not in the same synset.

Extensive accounts of cross-POS derivatives and intra-POS verbal derivatives are given in Maziarz et al. (2011) and Maziarz et al. (2012) for Polish WordNet 2.0. This approach significantly differs from those

---

[2] Pala and Hlaváčková (2007) stress that "due to a variety of relations that result from combinations of prefixes and base verbs (e.g., distributive, location, time, measure and others), this topic calls for a separate examination (project)".

mentioned above for other Slavic wordnets. Polish WordNet was not developed through the expand model, and its structure heavily relies on lexical units, i.e., literals (word-sense pairs). In Polish WordNet, aspectual pairs are kept apart and lexical hierarchies consist of either perfective or imperfective verbs. Relations between verbs are divided into purely semantic relations (inter-register synonymy, hyponymy and hypernymy, meronymy and holonymy, two types of antonymy, converseness, state, processuality, causality, inchoativity, presupposition, preceding, fuzzynymy) and derivationally-motivated relations (pure aspectuality, secondary aspectuality, iterativity, derivationality, cross-categorial synonymy, role inclusion). Some of the relations hold between lexical units (word-sense pairs, e.g., antonymy or pure aspectuality), while others hold between synsets (e.g., hyponymy and processuality). Although the relations in Polish WordNet do account for verbal derivatives, we think that a more fine-grained analysis is required for Slavic wordnets.

## 3 DERIVATIONAL RELATIONS BETWEEN CROATIAN VERBS

As in other Slavic languages, Croatian verbs are always marked for aspect and classified as perfective, imperfective, or bi-aspectual.[3] Imperfective and perfective verbs in Croatian can, in terms of derivation, be roughly divided into four groups. The first group comprises non-prefixed imperfective verbs (e.g., *pisati* 'to write'). The second group consists of predominantly perfective verbs built by prefixation of verbs from the first group (e.g., *pre+pisati* 'to copy by writing'). The third group comprises imperfective verbs that denote the iterativity of the action. Verbs in this group are built by suffixation of verbs from the second group (e.g., *pre+pis+iva+ti* 'to copy over and over again'). The fourth group consists of perfective verbs derived by prefixation of verbs from the third group (e.g., *is+pre+pis+iva+ti* 'to finish copying over and over again'). Verbs in this group include distributive verbs denoting actions performed by several agents usually on several ob-

---

[3] It is important to stress that so-called bi-aspectual verbs are not imperfective and perfective at the same time. They acquire a perfective or imperfective reading in a particular context.

jects.[4] Aspectual pairs are formed by prefixation or suffixation. Prefixation of an imperfective verb can either yield an imperfective verb (*osjećati* 'to feel' – *su* + *osjećati* 'to sympathize') or a perfective verb (*pisati* 'to write' – *pre* + *pisati* 'to copy'). Prefixation of a perfective verb can yield only perfective forms (*dati* 'to give' – *pre* + *dati* 'to hand over'). Imperfective forms of perfective verbs are built by suffixation (*dati* 'to give' – *da* + *va* + *ti* 'to give repeatedly'). Whereas both types of affixes can be applied to create pure aspectual pairs, only suffixation is used to generate iterative forms of perfective verbs. Aspectual pairs can thus be divided into primary or true aspectual pairs and secondary aspectual pairs. True aspectual pairs are determined primarily by the test of secondary imperfectivization (cf. Raguž, 1997; Jelaska et al., 2005; Maziarz et al., 2011), but also by additional criteria pertaining to the semantics of prefixes. The relation of pure aspectuality exists between a base form and a derivative with a prefix which does not contain any other semantic components except perfectiveness. For example, although the verb *potrčati* 'to start running' does not pass the test of secondary imperfectivization (it is not possible to derive an iterative imperfective \**potrčavati*), it is not the pure aspectual pair of the verb *trčati* 'to run$_{ipf}$' since it additionally denotes the beginning of the action. In other words, the lexical meaning of the verb *potrčati* contains the aspectual component of perfectiveness and the semantic component of inchoativity. The lexical meaning of the verb *dotrčati* 'to run to$_{pf}$', which denotes the other end of the same action, contains the aspectual component of perfectiveness as well as semantic components of completeness and direction of movement. Therefore, the true or primary aspectual pair of the verb *trčati* is the derived form *otrčati* 'to run$_{pf}$'[5]. This derived form consists of the prefix *od*- and the imperfective infinitive *trčati* 'to run$_{ipf}$'. It is a polysemous unit with two different meanings. The first meaning is locative, namely, 'to run from', while the second is 'to finish running'. Due to its second meaning, the derivative *otrčati* is the primary aspectual pair of

---

[4] Distributive verbs have several subjects (*poiskakati* 'to jump one by one'), several objects (*poubijati* 'to kill one by one') or both several subjects and objects (*pogledavati* 'to glance at each other').

[5] Change from *od* + *trčati* via *ot* + *trčati* to *otrčati* is morphonologically determined (voiced 'd' is devoiced in front of the voiceless 't' and then blended into a single 't').

the verb *trčati*, since the only difference in their meanings is that of aspect. The verb *otrčati* does not pass the test of secondary imperfectivization, and its prefix is regarded as semantically most deprived of its content in comparison to other derivatives. As a general rule we postulate that if a derivational prefix does not have additional semantic components (e.g., inchoativity) apart from the aspectual component of perfectiveness and the derivative does not pass the test of secondary imperfectivization, then the base form and its derivative are true or primary aspectual pairs. Verbs that do not pass the test of secondary imperfectivization, but whose prefixes simultaneously have additional semantic components are not considered true aspectual pairs. This distinction is important since true aspectual pairs are members of the same synset in CroWN. Although the relation between pure aspectual pairs, as well as between secondary aspectual pairs, is considered to be a derivational phenomenon in Croatian linguistics (cf. Barić et al., 2003; Babić, 2002), we treat them as members of same synsets in CroWN since the only difference in meaning between them is the difference in aspect. The same holds for iterative verbs and prefixed perfectives that serve as their base forms. Although iterative verbs have the additional semantic component of repetitiveness, they are grouped in the same synset as their base forms since their lexical meaning is not affected by this additional component. However, the difference in aspect is reflected in the definitions of synset meanings. Each synset member is tagged with one of the following aspect labels: IPF, PF, BI, or ITER, representing imperfective, perfective, bi-aspectual and iterative forms. This distinction is also reflected in different aspectual forms used in definitions that vary according to the aspect of the literals they relate to.[6]

In addition to aspect change, both prefixes and suffixes can create a shift in the meaning of base forms, but the semantic impact of suffixes is rather limited. Apart from the change in aspect, suffixes are used to form verbs denoting diminutive actions and pejorative attitudes. On the other hand, the semantic impact of prefixes is much wider and less predictable. Combinations of prefixes and base forms

---

[6] Definitions are structurally and semantically the same. The only difference is that imperfectives are defined with imperfective verbs; perfectives with perfectives; bi-aspectual verbs with combinations of imperfective and perfective verbs; and iteratives with imperfectives + 'repeatedly'.

can vary in terms of meaning from compositional to completely idiosyncratic.

## 4 PREFIXATION

Prefixation is the most productive derivational process of Croatian verbs (Babić, 2002). Verbs are derived by prefixation only from other verbs. There are 19 productive prefixes in Croatian: *do-, iz-, na-, nad-, o-/ob-, obez-, od-, po-, pod-, pre-, pred-, pri-, pro-, raz-, s-, su-, u-, uz-, za-*.[7] The majority of these prefixes are of prepositional origin and have homographic counterparts in prepositions. The prefixes without prepositional counterparts in contemporary Croatian are *obez-*[8], *pre-, pro-, raz-* and *su-*. Prefixation of Croatian verbs can trigger:

(a) a change in aspect (e.g., *puniti* 'to fill$_{ipf}$' – *napuniti* 'to fill$_{pf}$');

(b) a change in aspect and the addition of a new, more specific semantic component to the base form (e.g., *puniti* 'to fill$_{ipf}$' – *ispuniti* – 'to fill something completely$_{pf}$');

(c) the addition of a new component to the meaning of the base form without a change in its aspect (*osjećati* 'to feel$_{ipf}$'– *suosjećati* 'to sympathize$_{ipf}$').

Since prefixes are developed from prepositions, most of them retained their original prepositional meanings. The semantic structure of prefixes can be described as a radial polysemous structure with one central and several peripheral meanings.[9] Very often derivatives acquire two or more of their semantic components. One of them is always more prominent than the others that are nevertheless present in the overall

---

[7] Apart from these 19 prefixes, there are several non-productive prefixes, as well as prefixes of foreign origin not taken into consideration here.

[8] This prefix is actually a combination of two prepositions (*o* + *bez*), but *bez-* cannot apear as an individual prefix.

[9] Polysemous units as categories with radial structure consisting of a central and several peripheral meanings (determined by metonymical and metaphorical shifts) are presented by Lakoff (1987); Langacker (1987); Raffaelli (2007), and others. Prepositions as polysemous units are analyzed, for example, by Lakoff (1987), and Lindner (1981). For a Slavic analysis of Russian prefixes (cf. Janda, 1985, 1986). Croatian prepositions and prefixes are analysed in the cognitive framework by Šarić (2003, 2006a,b) and Belaj (2008a,b).

semantic structure of derivatives. The polysemous structure of Croatian prefixes can be illustrated with the prefix *za-*, which developed from the preposition *za*. This preposition has several locative meanings, as e.g., *behind* (*Metla je za ormarom* 'There is a broom behind the closet') or *at* (*Sjedi za stolom* 'He sits at the table'). This preposition can be used for expressing temporal relations, as e.g., *during* (*Za vrijeme mog boravka…* 'During my stay…') and *after* (*Dolazi za mnom* 'He will arrive after me'), but also in adverbial constructions of quantity or manner. The semantic complexity of the preposition *za* is also reflected in the polysemous structure of the prefix *za-*. As other prefixes, *za-* can be used for the derivation of pure aspectual pairs, but also for the production of derivatives with specific meanings, e.g.:

(a) locative meaning
   (1) to put something behind something (*zabaciti* 'to throw behind')
   (2) to put something onto something (*zakačiti* 'to attach')
   (3) to change position (*zaleći* 'to lie down')
   (4) to move around (*zakrenuti* 'to go in a curve or around the corner')
(b) inchoative meaning (*zapjevati* 'to start singing')
(c) more or less intensified action (*zamisliti se* 'to ponder', *zagorjeti* 'to scorch')
(d) change of property (*zacrvenjeti se* 'to become red')
(e) pure aspectuality (*zaklati* 'to slaughter')

Despite such complex semantic structure, we believe that the specific meanings of the prefix *za-*, apart from pure perfective meaning in (e), can be divided into four larger classes which we label: (1) location, (2) quantity, (3) time, and (4) manner. In a similar analysis applied to other productive prefixes mentioned above, we tried to deduct their meaning components and their impact on the meaning of the derived verbs. The resulting combinations of prefixes are divided into two groups: (1) pure aspectual pairs and (2) secondary aspectual pairs (cf. Figure 1). The first one is described above and we shall not go into further details.

The group of secondary aspectual pairs is further divided into two broader classes according to the semantic criterion: (1) compositional

Figure 1: Meanings of the prefix *za-*
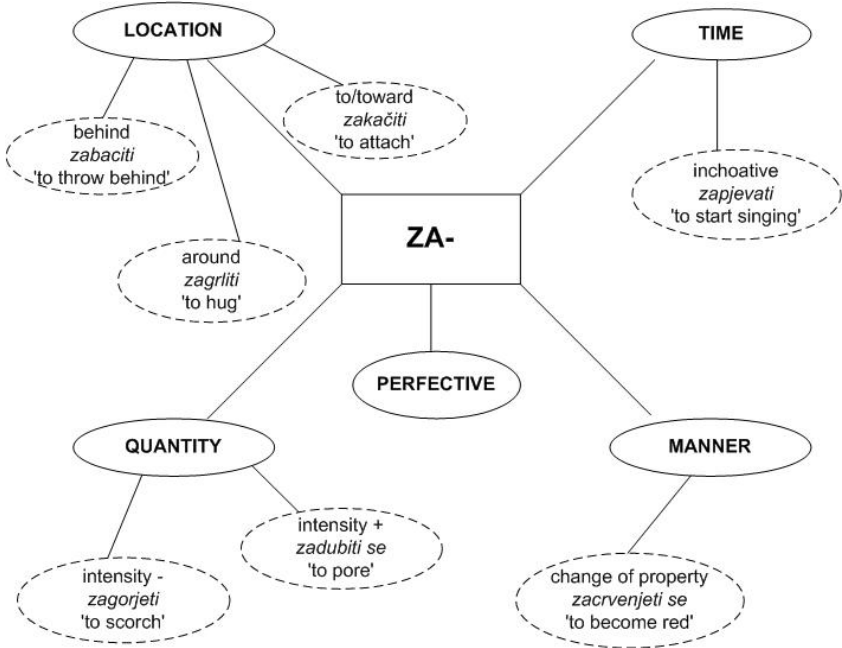


and (2) idiosyncratic. The division is motivated by the extent of the semantic shift that takes place in derived forms. Combinations of prefixes and base verbs in Croatian form a continuum in terms of semantic compositionality. On one pole of this continuum there are compositional combinations, i.e., one of the specific meanings of a prefix and lexical meaning of a verb are semantically transparent (e.g., *govoriti* 'to speak' – *progovoriti* 'to start speaking', *pjevati* 'to sing' – *zapjevati* 'to start singing'). On the other pole of this continuum there are completely idiosyncratic combinations. In these combinations the meaning of the derivatives as a whole cannot be directly connected either to the meaning of the prefix or to the lexical meaning of the verb without a thorough analysis of metaphorical or metonymical shifts (e.g., *baciti* 'to throw' – *pobaciti* 'to abort pregnancy'; *pustiti* 'to release' – *napustiti* 'to abandon').

In further sections we focus on predominantly compositional combinations. The goal of this research is to detect and describe meanings of prefixes that are constant and present in combinations with base verbs, i.e., those prefixal meanings that occur even when attached to
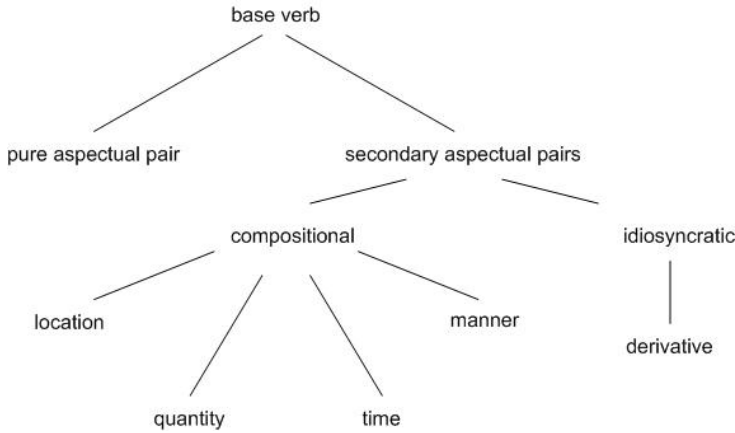
Figure 2:
Derivationally motivated
relations between base
verbs and derivatives

verbs from various semantic fields.[10] The objective of this procedure is first to determine the set of prefixal meanings that reoccur in various semantic fields and secondly to determine which prefixes can carry the same meanings. The final objective is to establish the set of derivationally motivated semantic relations between Croatian verbs. We will further refer to these relations as *morphosemantic relations*. These are further analyzed in order to determine which relations should be introduced into Croatian WordNet, since they are not encompassed by the existing semantic relations.

To fulfill these tasks, it is necessary to determine which prefixes take part in the derivation of particular base forms. The data on the derivational spans of verbs so far have not been systematically and extensively presented in Croatian morphology. In other words, large-scale data indicating which affixes are used or can be used with particular base forms in Croatian do not exist.

4.1                     *Derivational Database*

In order to address these issues, we have collected approximately 14,000 verbal lemmas from digital and freely available dictionaries of Croatian. The initial list consisted of infinitives unsorted in any way. The verbs from the list were automatically processed using a rule-based approach. In the first step of processing we applied a set of rules

---

[10] Verbs are divided into 15 semantic fields in PWN (cf. Fellbaum, 1998). The semantic fields were taken from WordNet 1.5 (so-called "lexicographic files") and mapped onto verbal synsets in CroWN.

for the segmentation of prefixes in order to obtain base forms and their derivatives formed through prefixation. The set of rules was designed to remove the 19 productive prefixes presented in the Section 4, as well as their combinations. In Croatian, one base form can bear one, two, three and very rarely even four derivational prefixes at the same time (e.g., $pre_{\text{prefix1}} + ras_{\text{prefix2}} + po_{\text{prefix3}} + dijeliti_{\text{base-form}}$ 'to reassign, to reallocate'). On the other hand, only one derivational suffix is added to roots. Besides a derivational suffix, stems can have either zero or one conjugational suffix before the infinitive endings *-ti* or *-ći*. Conjugational suffixes indicate verbal inflectional classes. The second set of rules was created to recognize and segment the suffixes and the roots. The verb *izrezuckati* 'to cut into small pieces$_{\text{pf}}$' was thus segmented into a prefix (*iz-*), a root (*-rez-*), a derivational suffix (*-uck-*), a conjugational suffix (*-a-*), and an infinitive ending (*-ti*). A form without any derivational affix, i.e. the base form of this derivative, is the verb *rezati*$_{\text{ipf}}$. The aim of this procedure was also to obtain a set of base forms that are either non-prefixed infinitives, i.e. lemmas, or morphological stems. These morphological stems are not lemmas, although they are used in further derivational processes. For example, the stem *\*laziti* can acquire different prefixes and thus serves as the base form for the derivation of verbs such as *do-laziti* 'to come$_{\text{ipf}}$', *iz-laziti* 'to exit$_{\text{ipf}}$', *pre-laziti* 'to cross$_{\text{ipf}}$' etc., but it cannot stand alone as an individual word.

In numerous cases, the rules could not detect an affix due to graphical overlapping with its stem. Prefixes were not accurately detached when they were graphically identical to parts of stems. For example, verbs like *privilegirati* 'to privilege' or *sniježiti* 'to snow' were incorrectly segmented as *\*pri + vileg + ir + a + ti* instead of *privileg + ir + a + ti* and *\*s + nijež + i + ti* instead of *snijež + i + ti*. A similar problem occurred with suffixes and roots. For example, *krijumčariti* 'to smuggle' was segmented into *krijum + čar + i + ti* instead of *krijumčar + ø + i + ti* and pobjeći 'to run away' into *po + b + ø + je + ći* instead of *po + bje + ø + ø + ći*. The output of processing was therefore manually checked and corrected. The final result of this rule-based semiautomatic procedure is a derivational database consisting of infinitives segmented into lexical and grammatical morphemes. This database has enabled further exploration of the derivational network of verbs sharing the same base form. A sample of the database is given in Figure 3. Each lexical entry in the derivational database consists
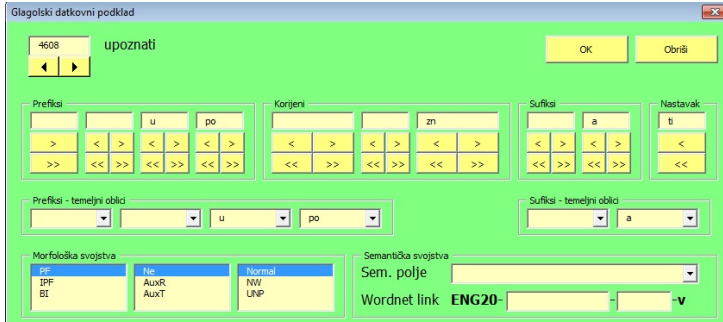
Figure 3:
Sample of derivational
database of Croatian verbs

of verbs decomposed into groups of morphemes, and each group of morphemes is provided with slots for roots, affixes and linguistic meta-data. Slots for morphemes are divided into: 1. derivational prefixes (four slots)[11], 2. the lexical part (three slots – in the majority of cases only one slot is filled, the three slots are provided for verbal compounds of two lexical morphemes and an interfix), 3. derivational and conjugational suffixes (two slots) 4. infinitive ending (one slot). The meta-data in lexical entries indicates verbal aspect, types of reflexivity, etc. The database enables queries across the full derivational span of particular base forms and generalizations regarding the distribution and frequency of affixes in the derivation of verbs from other verbs and verbal stems. In its present shape the database comprises 16,834 entries consisting of 14,291 lemmas and 2543 productive stems, i.e. the stems used the in formation of at least 2 derivatives.[12] In the remainder of the paper we focus on combinations of one prefix and a base form, the most productive derivational process among Croatian verbs according to the data from the database. The database has enabled the recognition of 4221 unprefixed base forms used in the prefixal derivation of 10,070 verbs. The distribution of prefixes across slots in the database is given in Figure 4 (P1 – the first slot next to the root contains a prefix, P2 – the first and the second slot next to the root contain prefixes etc.).

---

[11] Combinations of various prefixes differently affect the meaning of the base form. The combinations and derivations possible from these derived forms are the subject of further research. Preliminary results are shown in Šojat et al. (2012).

[12] The derivational database will be further expanded in terms of other parts of speech. Queries over the database will be possible through a web interface, which is still under construction.

Figure 4:
Distribution of
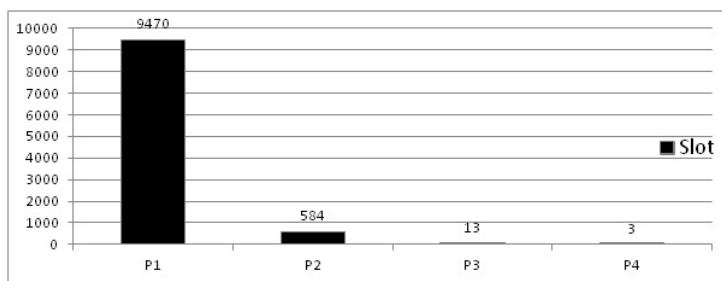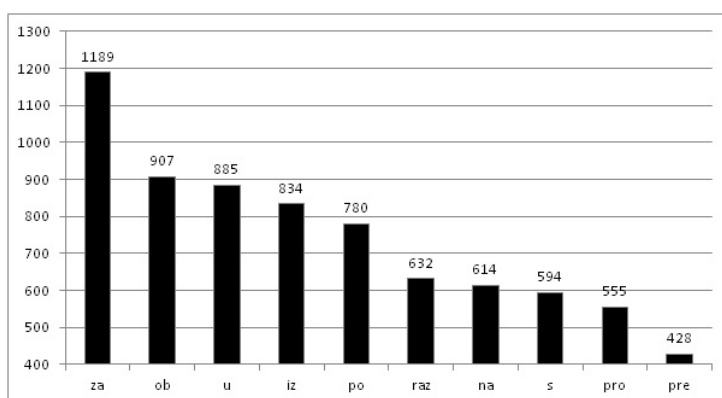prefixes across slots in
derivational database



Figure 5:
Frequency of particular
prefixes in the derivational
database: 10 most frequent
prefixes in the P1 slot



The ten most frequent prefixes in such combinations are given in Figure 5. The recognition of possible and attested combinations of prefixes and base forms has enabled the classification of prefixal meanings into broad and general categories mentioned in Section 4 above. These categories serve as a basis for further analysis and elaboration of the morphosemantic relations between verbal base forms and verbal derivatives.

## 4.2 *Prefixal Meanings*

Prefixes in Croatian usually have various and heterogeneous meanings which are often hard to capture and to separate from one another within the same prefix. As mentioned, we have focused on predominantly compositional combinations of prefixes and base forms. In order to establish the set of morphosemantic relations among verbs we have divided prefixal meanings into four major groups: (1) location, (2) time, (3) quantity and (4) manner. These four major groups were further divided into several subgroups. The division into four major

groups is built upon already existing categorizations of prefixal meanings in Croatian grammars (cf. Babić, 2002; Barić et al., 2003) and a preliminary systematization for the purposes of introducing morphosemantic relations into CroWN (cf. Srebačić, 2011). In this paper we have further divided the four major groups into several subgroups upon more in-depth analysis of the prefixal meanings, presented below.

In the *location* group, the prepositional origin of prefixes pervades in their meaning and is more prominent than in the other groups. In this group, prefixes primarily denote spatial relations, i.e., a particular direction or location. The *time* group includes prefixal meanings that refer to various phases of the action denoted by base verbs, such as the beginning or the termination of the denoted action. The *quantity* group includes prefixal meanings that refer to amount or intensity of the action as determined by a prefix. Finally, the *manner* group includes prefixal meanings related to various modes of action denoted by base verbs. Table 1 lists all 19 prefixes and all their meanings established in the analysis and used in further processing.

Such a thorough analysis of 19 productive prefixes in Croatian and their meanings enabled the establishment of major groups and subgroups of morphosemantic relations, which will be presented in the following section.

## 5        MORPHOSEMANTIC RELATIONS

Labels for morphosemantic relations consist of two parts. The first one pertains to one of the four major groups: location, time, quantity, and manner. The second part indicates particular subgroups of major groups.

**location_**

1. loc_bott_up – upward movement
2. loc_top_down – downward movement
3. loc_prox – movement in proximity to a subject or object
4. loc_through – movement through something (or someone)
5. loc_apart – movement in opposite or multiple directions
6. loc_to_toward – movement to or toward something or someone

Table 1: The meanings of verbal prefixes

| Prefix | Location | Time | Quantity | Manner |
|--------|----------|------|----------|--------|
| do- | 1. **to/toward** – *doletjeti* 'to fly to' | 1. **completion** – *dozreti* 'to ripen', *dostaviti* 'to deliver' 2. **finitiveness** – *domisliti se* 'to think out' | 1. **addition** – dodati 'to add', dopisati 'to add in writing' | |
| iz- | 1. **bottom-up** – *izrasti* 'to grow up', *izroniti* 'to emerge' 2. **from** – *izletjeti* 'to fly from', *izliti* 'to pour out' | 1. **distributivity** – *izbacati* 'to throw out one by one' , *ispisati* 'to print' 2. **completion** – *izliječiti* 'to cure' | 1. **sufficiency** – **isplakati se** 'to cry one's eyes out', *izvikati se* 'to shout one's fill' 2. **excessiveness** – *izmučiti se* 'to exhaust oneself' | |
| na- | 1. **top-down** – *nabosti* 'to prick, to spike' 2. **proximity** – *naići* 'to come across' 3. **to/toward** – *nalijepiti* 'to paste' | 1. **inchoativity** – *natrunuti* 'to begin to rot', *nagristi* 'to start to bite/corode' 2. **distributivity** – *navoziti* 'to cart one by one' | 1. **sufficiency** – *najesti se* 'to stuff oneself' 2. **excessiveness** – *napiti se* 'to get drunk' 3. **intensity** – *naraditi se, namučiti se* 'to tire oneself out with work', *nagorjeti* 'to scorch' 4. **addition** – *naloviti* 'catch a quantity of something' | |
| nad- | 1. **over** – *nadgraditi* 'to outbuild'; *nadletjeti* 'to fly over' | | 1.**exceeding** – *nadrasti* 'to outgrow', *nadjačati* 'to overpower' | |
| o-/ob- | 1. **around** – *okružiti* 'to encircle'; *oploviti* 'to circumnavigate, to sail around' | | | |

| | | | | |
|---|---|---|---|---|
| obez- | | | 1. **deprivation** – *obezvrijediti* 'to devaluate' | |
| od- | 1. **apart** – *odletjeti* 'to fly away', otići 'to leave' | 1. **completion** – *odigrati* 'to play', *odsvirati* 'to play a musical piece' | | |
| po- | 1. **top-down** – *poleći* 'to lay down', *posoliti* 'to salt' | 1. **inchoativity** – *potrčati* 'to start running', *poletjeti* 'to start flying' 2. **distributivity** – *pomrijeti* 'to die one by one', *pobiti* 'to kill one by one' | 1. **intensity** – *poprati* 'to wash a little', *poigrati se* 'to play a little' | |
| pod- | 1. **under** – *podbosti* 'to spur', *podložiti* 'to place under' | | 1. **insufficiency** – *potplatiti* 'to underpay', *pothraniti* 'to feed insufficiently' | |
| pre- | 1. **over** – *preskočiti* 'to jump over', *preletjeti* 'to fly over' 2. **re-location** – *preseliti* 'to relocate', *pretočiti* 'to pour over' | 1. **completion** – *prenoćiti* 'to spend the night' | 1. **intensity** – *presoliti* 'to oversalt', *pregrijati* 'to overheat' 2. **exceeding** – *prerasti* 'to outgrow' | 1. **change of property** – *pretvoriti se* 'to convert', *preimenovati* 'to rename' |
| pred- | | 1. **preceding** – *pretplatiti se* 'to pay in advance', *prethoditi* 'to precede' | | |
| pri- | 1. **proximity** – *primaknuti se* 'to come closer', 2. **to/toward** – *prikačiti* 'to attach', *pribiti* 'to pin down' | | 1. **intensity** – *primiriti se* 'to calm down a little' 2. **addition** – *priliti* 'to add by pouring' | 1. **connection** – *prišiti* 'to sew on' |
| pro- | 1. **through** – *probiti* 'to break through', 2. **proximity** – *projuriti* 'to pass quickly by', *prohujati* 'to rush by' | 1. **inchoativity** – *progovoriti* 'to start talking' 2. **completion** – *prožvakati* 'to finish chewing' 3. **preceding** – *proreći* 'to predict' | 1. **intensity** – *prodrmati* 'to shake a little', *proprati* 'to rinse' | |

| | | | | |
|---|---|---|---|---|
| raz- | 1. **apart** – *razdvojiti se* 'to separate', *raširiti se* 'to spread' | | 1. **intensity** – *razljutiti se* 'to become very angry' | |
| s- | 1. **top-down** – *srušiti* 'to knock down, to fell', *sletjeti* 'to land' | | | 1. **connection** – *spojiti* 'to bond, to bring together' |
| su- | 1. **proximity** – *susresti se* 'to meet', *sudariti se* 'to bump' | | | 1. **connection** – *sufinancirati* 'to cofinance' 2. **opposition** – *sučeliti se* 'to face' |
| u- | 1. **into** – *uplivati* 'to swim into', *urasti* 'to grow into' | 1. **finitiveness** – *ugaziti* 'to trample' | 1. **intensity** – *usjedjeti se* 'to sit for a long time', *uznojiti se* 'to sweat abundantly' | 1. **change of property** – *usmrdjeti se* 'to become stinky', *uprljati se* 'to become dirty' |
| uz- | 1. **proximity** – *uspinjati se* 'to climb', *uzdizati se* 'to ascend' | 1. **inchoativity** – *uskomešati se* 'to stir up' | 1. **intensity** – *uzburkati* 'to stir up', *ushodati se* 'to walk up and down' | |
| za- | 1. **around** – *zagrliti* 'to hug' 2. **behind** – *zabaciti* 'to throw back' 3. **to/toward** – *zakačiti* 'to attach' 4. **top-down** – *zaleći* 'to lie down' | 1. **inchoativity** – *zatrčati se* 'to start running', *zapjevati* 'to start singing' | 1. **intensity** – *zadubiti se* 'to pore', *zagorjeti* 'to scorch' | 1. **change of property** – *zacrveniti se* 'to become red' |

7. loc_over – movement over something or someone
8. loc_into – movement into something (or someone)
9. loc_around – movement around something or someone
10. loc_under – movement or location beneath something or someone
11. loc_reloc – movement to another location
12. loc_behind – movement behind something or someone
13. loc_across – movement across something
14. loc_from – movement away from something or someone

This group predominantly consists of verbs of movement, since various spatial relations are inherent in their lexical meanings. These

| Location | Prefix |
|----------|--------|
| **bottom-up** – *uspeti se* 'to climb', *izrasti* 'to grow up' | iz-, po-, uz- |
| **top-down** – *porušiti* 'to pull down' , *nabosti* 'to spike', *sletjeti* 'to land' | na-, po-, s-, za- |
| **proximity** – *naići* 'to come across', *približiti se* 'to come closer', *projuriti* | na-, pri-, pro-, su- |
| **through** – *probiti* 'to break through', *prošiti* 'to quilt' | pro- |
| **apart** – *odvojiti* 'to separate', *otkinuti* 'to detach' | od-, raz- |
| **to/towards** – *prikačiti* 'to attach', *zabiti* 'to nail', *nalijepiti* 'to stick' | na-, pri-, za- |
| **over** – *natkriti* 'to cover over', *preskočiti* 'to jump over' | nad-, pre- |
| **into** – *utrčati* 'to run into', *urasti* 'to grow into' | u- |
| **around** – *okružiti* 'to circle', *obletjeti* 'to fly around something', *obuhvatiti* 'to embrace' | o-/ob-, za- |
| **under** – *podrediti* 'to subject', *podložiti* 'to place under' | pod- |
| **re-location** – *preliti* 'to decant', *preseliti* 'to move' | pre- |
| **behind** – *zabaciti* 'to throw back' | uz-, za- |
| **across** – *prijeći* 'to cross', *preletjeti* 'to fly over', *preplivati* 'to swim across' | pre- |
| **from** – *izletjeti* 'to fly from', *izliti* 'to pour out' | iz- |

Table 2: Morphosemantic relations in *location* group

relations also hold between numerous base verbs and their derivatives from other semantic fields, e.g., *prošiti* 'to quilt', *preliti* 'to pour over'. Due to their prepositional origin, prefixes primarily denote spatial relations. For this reason, the majority of prefixes have at least one meaning corresponding to one of the location relations. This fact in turn results in a rather extensive set of morphosemantic relations of location. All location morphosemantic relations with examples are listed in Table 2.

**time_**

1. time_inch – beginning of the action ('to start X'[13])
2. time_fin – termination of the action ('to finish X')

---

[13] X = base verb.

Table 3:
Morphosemantic
relations in *time*
group

| Time | Prefix |
|---|---|
| **inchoativity** – *pojuriti* 'to start rushing', *zaplivati* 'to start swimming', *prozboriti* 'to start talking' | na-, po-, pro-, uz-, za- |
| **finitiveness** – *doletjeti* 'to fly to', *dotrčati* 'to run to' | do-, na-, u- |
| **distributivity** – *izdijeliti* 'to give one by one', *popadati* 'to fall one by one' | iz-, na-, po- |
| **preceding** – *pretkazati* 'to predict', *prethoditi* 'to forego', *pretplatiti* 'to subscribe' | na-, pred-, pro- |

3. time_distr – the action performed by several subjects usually on several objects and in successive phases ('repeatedly X')

4. time_prec – the action denoted by derivatives precedes the action denoted by base verbs[14]

The group of time relations is determined by aspectual properties and constraints of Croatian verbs. Relations in this group do not hold between pure aspectual pairs. Besides the aspectual difference between imperfective base forms and perfective derivatives, derivatives also denote various phases or temporal components of the denoted action, such as its starting or terminative point. Morphosemantic relations belonging to the time group with examples are in Table 3.

**quan_**

1. quan_suff – the action denoted by the derivative is performed in sufficient or insufficient quantity ('enough/not enough X')

2. quan_exc – the action denoted by the derivative is performed in excessive quantity ('too much X')

3. quan_int – the action denoted by the derivative is performed with weaker or stronger intensity ('X a little/a lot')

4. quan_more – the action denoted by the derivative outperforms the action denoted by the base verb that is performed by one or more different subjects ('X better than')

---

[14] The pure semantic relation *preceding* exists in Polish WordNet (cf. Maziarz et al., 2011), where this relation is used between synsets. In our approach, this relation holds between derivationally related verbs.

| Quantity | Prefix |
|---|---|
| **sufficiency** ( + / − ) – *istrčati se* 'to run enough'; *potplatiti* 'to underpay', *pothraniti* 'to feed insufficiently' | iz-, na-, pod- |
| **excessiveness** – *napiti se* 'to get drunk', *prejesti se* 'to gormandize', *izmučiti se* 'to exhaust oneself' | iz-, na-, pre- |
| **intensitiy** ( + / − ) – *nagristi* 'to bite a little', *protresti* 'to shake a little', *ustrčati se* 'to bustle around', *razbjesniti se* 'to become very furious' | na-, po-, pre-, pri-, raz-, u-, uz-, za- |
| **exceeding** – *nadigrati* 'to outplay', *nadrasti* 'to outgrow' | nad-, pre- |
| **deprivation** – *obeshrabriti* 'to discourage', *obezbojiti* 'to decolour' | obez- |
| **addition** – *dogrijati* 'to heat to the desirable degree', *dopisati* 'to add by writing' | do-, na- |

Table 4: Morphosemantic relations in *quantity* group

5. quan_depr – the action denoted by the derivative refers to the loss of property[15]

6. quan_add – the action denoted by the derivative refers to the addition or completion of the action denoted by the base verb ('to add by X-ing')

As far as we know, quantity as a morphosemantic category has not been accounted for in related work or lexical resources. Our analysis has shown, however, that it must be taken into consideration when dealing with the prefixation of Croatian verbs and the morphosemantic relations between base forms and derivatives. Moreover, since it comprises six subgroups, we firmly believe this group is well justified. Morphosemantic quantitative relations with examples are listed in Table 4.

**mann_**

1. mann_conn – the action denoted by the derivative refers to two or more inter-related entities. This relation comprises *connection* and *opposition* as stated in Table 1.

---

[15] Although properties are generally expressed by adjectives, there are verbs derived from adjectives denoting the same property. This relation holds between such verbs and their verbal derivatives (e.g., *obeshrabriti* 'to discourage' is derived from the base verb *hrabriti* 'to encourage', which is in turn derived from the adjective *hrabar* 'courageous'.

| Manner | Prefix |
|---|---|
| **inter-connection** – *sudjelovati* 'to co-participate', *prikrpati* 'to sew on by patching', *sjediniti* 'to compound' | pri-, s-, su- |
| **change of property** – *zazelenjeti se* 'to become green', *ukiseliti se* 'to become sour' | o-/ob-, po-, pre-, s-, u-, za- |

2. mann_prop – the action denoted by the derivative refers to the acquisition of property denoted by the base verb

The *manner* group consists of only two subgroups, but these specific meanings cannot be subsumed by any other major groups of relations. Each subgroup of relations is expressed by three or more different prefixes, forming a rather coherent and delimited group of meaning components. *Manner* morphosemantic relations with examples are in Table 5.

We also came across numerous derivatives that cannot be directly connected to their base verbs via any of the listed morphosemantic relations. As mentioned, there are combinations of prefixes and base verbs that are completely idiosyncratic. We mark the relation between such verbs with the underspecified relation *derivative* (cf. Maziarz et al., 2011).

### 5.1 *Morphosemantic relations in CroWN*

As mentioned above, our final objective was to establish the set of morphosemantic relations between Croatian verbs and determine which relations should be introduced into Croatian WordNet since they are not encompassed by the existing semantic relations.

CroWN contains 2318 verbal synsets with an average of 5.8 verbs per synset. Each verbal synset consists of verbs marked for their senses (so-called *literals*). The total number of verb senses, i.e. literals, is 13,476.[16] For example, *dati* ' to give$_{ipf}$' is marked for 28 senses and thus appears in 28 different synsets and *letjeti* 'to fly$_{ipf}$' is marked for 9 senses and appears in 9 different synsets.[17] There are 13 derivatives of

---

[16] PWN 3.0. comprises 25,047 verbal literals divided into 13,767 verbal synsets. Although the number of synsets in CroWN may seem rather small in comparison to the number of verbal synsets in PWN, the number of verb literals in CroWN is ca. 50% of the number of verb literals in PWN.

[17] Such a particularization of meaning is a consequence of the adopted expand

the verb *letjeti* formed with 8 different prefixes. In only one case does it occur that this base form and a derivative are aspectual pairs and therefore members of the same synset. The remaining 8 base forms and 12 derivatives are members of different synsets, and in more than 50% of cases they are members of different lexical hierarchies based on the semantic relation of hyponymy. In other words, *letjeti* 'to fly$_{ipf}$' is not positioned in the same hierarchy as the verb *uletjeti* 'to fly into$_{pf}$', even though they differ only in the meaning component 'moving into something'. In CroWN we use the same semantic relations between verbal synsets as in EWN and BN. These relations are synonymy, hyponymy/hyperonymy, antonymy, cause, and subevent. The relation of hypernymy/hyponymy is the most important for the overall structure of the lexicon. This relation can be described as 'to do X in a particular manner', where X is a hypernym. For example, verbs of movement are divided into several subfields on the basis of their specific meaning properties, such as manner of movement, direction of movement, the medium in which the movement is performed, means of movement, etc. Such a division results in hierarchies containing heterogeneous groups of hyponyms connected to their co-hypernym only through this specific meaning component. For example, 'to move' has hyponyms in the subfield of direction such as 'to move upwards', 'to move downwards', 'to move across something', 'to move through something', 'to move over something', and 'to move around something'. These hyponymy subclasses contain verbs denoting different media of movement, vehicles, manner, speed, etc. Apart from the general similarity that pertains to the concept of movement, verbs in these subclasses have significantly different meanings and frequently share only one meaning component, e.g., 'to move into something' (*uplivati* 'to swim into$_{pf}$', *utrčati* 'to run into$_{pf}$', *uletjeti* 'to fly into$_{pf}$') or 'to start moving' (*potrčat*i 'to start running$_{pf}$', *zaplivati* 'to start swimming$_{pf}$', *poletjeti* 'to start flying$_{pf}$'). This in turn results in hierarchies that do not contain derivationally related verbs that sometimes differ only in this particular meaning component. Therefore, we have proposed a set of relations between derivationally related verbs which are usually scattered across different hierarchies. In order to determine which morphose-

---

model (cf. Section 1) and in many cases does not truly reflect semantic structure and relations between Croatian verbs.

mantic relations could or should be introduced between base forms and derivatives in CroWN we conducted an experiment consisting of several steps.

In the first step we removed the sense tags from the literals and reduced this list to single appearances of forms. In other words, literals as tokens were treated as morphological types, resulting in 5747 unique forms. This list was automatically filtered for those verbs containing combinations of 2 and 3 prefixes, verbs with derivational suffixes, and iterative verbs formed by conjugational suffixes (cf. Section 3). The filtering was done by matching this list with the data from the derivational database (cf. Section 4). The output was a list of 2530 base forms and derivatives with only one prefix. This list was further filtered for 754 derivatives marked as aspectual pairs in CroWN. Finally, we obtained a list of 1922 verbal types in CroWN. These forms were used in the second step of the experiment. In this step we segmented prefixed forms into prefixes and base forms, again matching them with the derivational database. Thus we obtained 572 base forms and 1350 derivatives as candidates for the assignment of established morphosemantic relations (cf. Section 5). In the final step, we automatically assigned morphosemantic relations, according to particular prefixes as listed in tables 2-5, to each derivative and manually checked the results. In this analysis we either: (1) eliminated all suggested relations when none of them was appropriate due to the idiosyncratic nature of the combinations and tagged them as DERIV (cf. Figure 2) or (2) we chose the appropriate relation from the total of suggested relations. The result of the whole procedure is a list of 572 base forms and 1204 prefixed verbs marked for morphosemantic relations as described above. The distribution of morphosemantic relations according to particular prefixes and their overall frequency is given in Table 6.

The overall statistics concerning the four major groups of morphosemantic relations and their subgroups, as well as the number of occurrences between base forms and derivatives from CroWN is given in Tables 7, 8, 9[18] and 10.

---

[18] Two morphosemantic relations marked by * in the quantity table do not occur between verbs in CroWN, although they do occur between verbs in the derivational database. This is due to the significantly smaller number of base formes and derivatives in CroWN than in the derivational database.

| Prefix (overall freq.) | Morphosemantic relation | Freq. |
|---|---|---|
| do- (35) | loc_to_toward | 21 |
| | time_fin | 11 |
| | quan_add | 3 |
| iz- (133) | loc_from | 62 |
| | time_fin | 42 |
| | quan_exc | 10 |
| | time_distr | 10 |
| | quan_suff | 5 |
| | loc_bott_up | 44 |
| na- (71) | loc_top_down | 12 |
| | quan_int | 12 |
| | quan_add | 11 |
| | loc_to_toward | 8 |
| | time_inch | 7 |
| | quan_exc | 5 |
| | time_distr | 2 |
| | time_fin | 2 |
| | loc_across | 1 |
| | loc_prox | 1 |
| | time_prec | 1 |
| o-/ob- (83) | loc_around | 71 |
| | mann_prop | 12 |
| od- (88) | loc_apart | 75 |
| | time_fin | 13 |
| po- (108) | quan_int | 41 |
| | loc_bot_up | 25 |
| | time_inch | 21 |
| | time_distr | 16 |
| | loc_top_down | 3 |
| | mann_prop | 2 |
| pod- (10) | loc_under | 5 |
| | quan_suff | 5 |
| pre- (61) | loc_over | 26 |
| | mann_prop | 14 |
| | quan_exc | 11 |
| | loc_reloc | 5 |
| | quan_int | 2 |
| | time_fin | 2 |
| | loc_across | 1 |

Table 6:
Prefixes and morphosemantic links

| Prefix (overall freq.) | Morphosemantic relation | Freq. |
|---|---|---|
| pred- (4) | time_prec | 4 |
| pri- (66) | loc_to_toward | 34 |
| | quan_int | 12 |
| | loc_prox | 10 |
| | mann_conn | 5 |
| | quan_add | 5 |
| pro- (81) | loc_through | 34 |
| | time_fin | 20 |
| | quan_int | 11 |
| | time_inch | 11 |
| | loc_prox | 3 |
| | time_prec | 2 |
| s- (59) | mann_conn | 27 |
| | loc_top_down | 26 |
| | mann_prop | 6 |
| su- (6) | mann_conn | 4 |
| | loc_prox | 2 |
| u- (61) | loc_into | 61 |
| | time_fin | 40 |
| | mann_prop | 28 |
| | quan_int | 3 |
| uz- (32) | quan_int | 8 |
| | loc_prox | 6 |
| | time_inch | 6 |
| | loc_behind | 2 |
| za- (140) | time_inch | 68 |
| | mann_prop | 27 |
| | quan_int | 16 |
| | loc_around | 15 |
| | loc_to_toward | 7 |
| | loc_top_down | 4 |
| | loc_behind | 1 |

| Group (overall freq.) | Subgroup | Freq. |
|---|---|---|
| | loc_apart | **141** |
| | loc_araond | 87 |
| | loc_from | 70 |
| | loc_to_toward | 70 |
| | loc_top_down | 68 |
| | loc_into | 60 |
| | loc_through | 34 |
| LOCATION (600) | loc_over | 24 |
| | loc_prox | 23 |
| | loc_bott_up | 6 |
| | loc_under | 6 |
| | loc_behind | 5 |
| | loc_reloc | 4 |
| | loc_across | 2 |

Table 7:
Frequency of morphosemantic links – *location* group

| Group (overall freq.) | Subgroup | Freq. |
|---|---|---|
| | time_fin | **132** |
| TIME (276) | time_inch | 109 |
| | time_distr | 28 |
| | time_prec | 7 |

Table 8:
Frequency of morphosemantic links – *time* group

| Group (overall freq.) | Subgroup | Freq. |
|---|---|---|
| | quan_int | **126** |
| | quan_exc | 25 |
| | quan_suff | 20 |
| QUANTITY (190) | quan_add | 19 |
| | quan_depr | 0* |
| | quan_more | 0* |

Table 9:
Frequency of morphosemantic links – *quantity* group

| Group (overall freq.) | Subgroup | Freq. |
|---|---|---|
| MANNER (122) | mann_prop | **88** |
| | mann_conn | 34 |

Table 10:
Frequency of morphosemantic links – *manner* group

6                    DISCUSSION

As mentioned in Section 5.1, the semantic relations between verbal synsets in CroWN are hypernymy/hyponymy, synonymy, antonymy, cause, and subevent. All of them hold between whole synsets and none of them holds between particular literals, i.e., base verbs and their derivatives. Sometimes base verbs and derivatives are connected via one of the semantic relations that holds between synsets. In these cases, base verbs and derivatives are members of different synsets. However, in the majority of cases, morphosemantic and semantic relations do not overlap. Moreover, none of our morphosemantic relations can be completely subsumed by any of these semantic relations in terms of their semantic content.

As far as the semantic relation of hypernymy/hyponymy is concerned, we have indicated that base verbs and their derivatives are often not members of same lexical hierarchies and thus close semantic relations resulting from derivational processes are not recognizable between them.

Antonymy exists between two derivatives of the same base (e.g., *doći* 'to arrive' – *otići* 'to leave'), but this relation does not exist between derivatives and a base form (*ići* 'to go' – *doći* 'to arrive' and *ići* 'to go' – *otići* 'to leave'). The semantic relation cause holds between synsets and denotes the relation between two actions, the first denoting the cause and the second denoting the result or the consequence of action denoted by the first verb (e.g., *hraniti* 'to feed' – *jesti* 'to eat' or *ciljati* 'to aim' – *pogoditi* 'to shoot'). Although the relation cause semantically partially overlaps with our morphosemantic relation change of property, cause can only encompass pairs such as *kiseliti* 'to pickle$_{ipf}$' – *ukiseliti* 'to pickle$_{pf}$', but not their reflexive counterparts denoting the non-agentive action, e.g., *kiseliti se* 'to become sour$_{ipf}$' – *ukiseliti se* 'to become sour$_{pf}$'. The relation of subevent in EWN and BN denotes the relation between two synsets referring to two simultaneous actions or to an action which is a part of the action denoted by another synset (e.g., 'to eat' has subevents 'to chew' and 'to swallow'). This relation does not refer to derivationally related literals and does not reflect particular parts of events, e.g., its beginning or terminating point, as our morphosemantic relations of inchoativity or finitiveness do.

Since we have marked all of the verbs from CroWN that have one prefix (approx. 30% of all verbs in CroWN) with morphosemantic or aspectual relations, and only two of our morphosemantic relations were not applied at all, we believe that our inventory of relations is well justified and applicable not only in CroWN, but also in word-nets for other Slavic languages. From the related work on other Slavic languages presented in Sections 2 and 4 (especially Footnote 7), it is clear that the same interplay between base verbs and derivatives regarding the semantic impact of prefixes can be found in all branches of Slavic languages. We are convinced that this is always the case in South Slavic languages and that it also holds for Czech, Polish, and Russian.

The problem of including derivational relations in Slavic word-nets is already recognized and discussed, as shown in Section 2. However, the solutions presented do not seem to be fine-grained enough to include all morphosemantic relations between verbal derivatives. Based upon our analysis of prefixal meanings and their classification, we believe that the notion of secondary aspectuality can be further analyzed and divided into at least four major subgroups: *time*, *location*, *quantity* and *manner*. We strongly believe that these four major groups of secondary aspectuality, due to the similarity of Slavic languages, can be applied to other Slavic wordnets without significant changes. The morphosemantic subrelations presented here are probably more language-specific, and their existence or possible implementation should be examined for each Slavic language, but we believe that most of them can be applied to other Slavic word-nets.

## 7          CONCLUSION

Since all the relations in Croatian WordNet hold between synsets and not between single verbs, so far it has not been possible to account for morphosemantic relations between base forms and their derivatives as described above. The same problem has been detected for other Slavic wordnets, but the presented solutions are not fine-grained enough.

The work done on the derivational database of Croatian verbs has enabled the restructuring of the relations between verbs in CroWN, their adaptation to the lexical properties of the Croatian language and

Figure 6:
Base form
*letjeti* 'to fly'
and its derivatives
with meanings and
morphosemantic
relations
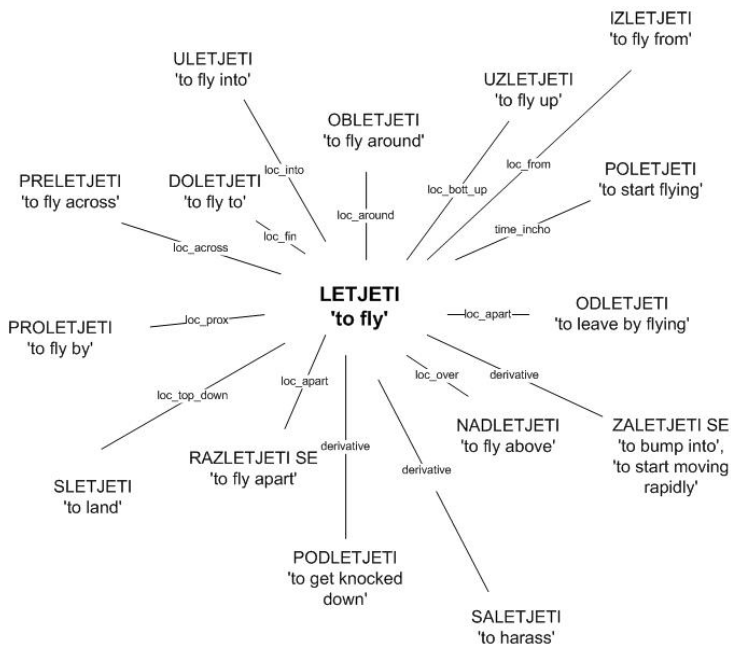


the enrichment of CroWN with morphosemantic relations as presented above.

The morphosemantic relations discussed here, resulting from combinations of one prefix and base forms, were first divided into four major groups and further into several subgroups. Combinations of multiple prefixes with the same base form and their influence on lexical meaning have yet to be investigated. This could potentially lead to a further expansion of the morphosemantic relations as stated here.

## REFERENCES

Stjepan Babić (2002), *Tvorba riječi u hrvatskome književnom jeziku*, Zagreb: HAZU: Nakladni zavod Globus.

Eugenija Barić et al.(2003), *Hrvatska gramatika*, Zagreb: Školska knjiga.

Branimir Belaj (2008a), *Jezik, prostor i konceptualizacija. Shematična značenja hrvatskih glagolskih prefiksa*, Osijek: Filozofski fakultet.

Branimir Belaj (2008b), Pre-locativity as the schematic meaning of the Croatian verbal prefix pred-, *Jezikoslovlje*, 9(1-2):123–140.

Orhan BILGIN, Özlem ÇETINOGLU and Kemal OFLAZER (2004), Building a WordNet for Turkish, *Romanian Journal of Information Science and Technology*, 7(1-2):163–172.

Bernard COMRIE (1989), *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*, Cambridge: Cambridge University Press.

Östen DAHL (1985), *Tense and aspect systems*, Oxford: Basil Blackwell.

Christiane FELLBAUM ed. (1998), *WordNet: An Electronic Lexical Database*, Cambridge: MA: MIT Press.

Christiane FELLBAUM, Anne OSHERSON and Peter E. CLARK (2007), Putting Semantics into WordNet's "Morphosemantic" Links, in *Proceedings of the Third Language and Technology Conference*, Poznan (Poland).

Laura JANDA (1985), The meaning of Russian Verbal Prefixes: Semantics and Grammar, Flier, A. S., Timberlake, A., eds. *The scope of Slavic aspect (UCLA Slavic Studies)*, Columbus, Ohio: Slavica, 26–40.

Laura JANDA (1986), *A Semantic Analysis of the Russian Verbal Prefixes ZA-, PERE-, DO- and OT-*, München: Otto Sagner.

Laura JANDA (2004), A metaphor in search of a source domain: the categories of Slavic aspect, *Cognitive Linguistics*, 15(4):471–527.

Zrinka JELASKA (2005), *Hrvatski kao drugi i strani jezik*, Zagreb: Hrvatska sveučilišna naklada.

Svetla KOEVA (2008a), Derivational and Morphosemantic Relations in Bulgarian WordNet, in *Proceedings of the Intelligent Information Systems 2008*, 359–368.

Svetla KOEVA, Cvetana KRSTEV and Duško VITAS (2008), Morpho-semantic Relations in WordNet – a Case Study for two Slavic Languages, in *Proceedings of the 4th Global WordNet Conference*, 239–254.

George LAKOFF (1987), *Women, Fire and Dangerous Things. What Categories Reveal about the Mind*, Chicago&London: The University of Chicago Press.

Ronald W. LANGACKER (1987), *Foundations of Cognitive Grammar. Vol. 1*, Stanford: Stanford University Press.

Susan J. LINDNER (1981), *A Lexico-Semantic Analysis of English Verb-Particle Constructions with UP and OUT*, PhD's dissertation, University of California, San Diego.

Marek MAZIARZ et al.(2011), Semantic Relations between Verbs in Polish WordNet 2.0, *Cognitive studies*, 11:183–200.

Marek MAZIARZ, Maciej PIASECKI and Stan SZPAKOWICZ (2012), An Implementation of a System of Verb Relations in plWordNet 2.0, in *Proceedings of the 6th Global WordNet Conference*, 181–188.

Karel Pala and Dana Hlaváčková (2007), Derivational Relations in Czech WordNet, in *Proceedings of the Workshop on Balto-Slavonic Languages*, 75–81.

Maciej Piasecki, Stan Szpakowicz and Bartosz Broda (2009), *A Wordnet from the Ground Up*, Wrocław University of Technology Press.

Ida Raffaelli (2007), Neka načela ustroja polisemnih leksema, *Filologija*, 48:135–172.

Ida Raffaelli et al.(2008), Building Croatian WordNet, in *Proceedings of the 4th Global WordNet Conference*, 349–360.

Dragutin Raguž (1997), *Praktična hrvatska gramatika*, Zagreb: Medicinska naklada.

Borislav Rizov (2008), Hydra: A Modal Logic Tool for Wordnet Development, Validation and Exploration, in *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, 1523–1528.

Ljiljana Šarić (2003), Prepositional categories and prototypes: Contrasting some Russian, Slovenian, Croatian and Polish examples, *Jezikoslovlje*, 4(2):187–204.

Ljiljana Šarić (2006a), A preliminary semantic analysis of the Croatian preposition u and its Slavic equivalents, *Jezikoslovlje*, 7(1-2):1–43.

Ljiljana Šarić (2006b), On the meaning and prototype of the preposition pri and the locative case: A comparative study of Slavic usage with emphasis on Croatian, *Rasprave instituta za hrvatski jezik i jezikoslovlje*, 32:225–248.

Krešimir Šojat, Nives Mikelić-Preradović and Marko Tadić (2012), Generation of Verbal Stems in Derivationally Rich Language, in *Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC'12)*, 928–933.

Matea Srebačić (2011), *Morphosemantic description of verbs of change in CroWN*, MA thesis, Department of Linguistics, Faculty of Humanities and Social Sciences, University of Zagreb.

Piek Vossen ed. (1998), *EuroWordNet. A Multilingual Database with Lexical Semantic Networks*, Dordrecht: Boston: London: Kluwer Academic Publishers.