

Morphosemantic fields in the building of the Croatian WordNet: the verbs of movement

Daniela Katunar and Krešimir Šojat

1. Introduction

The Croatian WordNet (henceforth CroWN) is a lexical-semantic database structured according to the principles of the Princeton WordNet (henceforth PWN). The basic unit of the PWN is a synset, i.e. a group of synonymous lexemes (literals) of the same part of speech (nouns, verbs, adjectives and adverbs). Each synset is linked with others via so-called semantic relations such as hyponymy, meronymy, antonymy and entailment (cf. Fellbaum 1998).¹ Pala and Hlaváčková (2007) stress that the semantic relations like synonymy, antonymy, hyponymy, meronymy as well as the additional relations introduced by the EWN, cannot capture all of the semantic relations in Czech, as well as in other Slavic languages involved. This claim is primarily aimed at various semantic relations that exist between morphologically interrelated lexemes from different synsets and often belong to different parts of speech. Pala and Hlaváčková (2007) also point out that the semantics of the derivational relations was not systematically analyzed and labeled either in the EWN or in the BN project.

The difference between English on the one hand and morphologically rich Slavic languages on the other has induced the enrichment of WordNets with new information primarily related to the inflectional and derivational patterns of the latter languages. Assuming that the derivational relation between two lexemes from different synsets is, in most cases, an indicator of mutual semantic relation, Koeva et al. (2008) point out that information related to inflectional properties has been added to all lexemes in the Bulgarian and Serbian WordNets, as well as some basic semantic relations that can be inferred from the derivational patterns, e.g. for possessive adjectives and for gender motion. Pala and Smrž (2004) and Pala and Hlaváčková (2007) present in their work a more extensive and elaborate attempt to semantically label the derivational relations that appear on a regular basis between lexemes, and include them in the Czech WordNet. To illustrate their regularity and validity for different roots and semantic fields Pala and Hlaváčková (2007: 77) use the Czech root(s) *-prác/prac-* ‘work’ and estab-

1 Most of the WordNets developed for other languages are based on the same or very similar structure, e.g. WordNets developed within the multilingual projects EuroWordNet and BalkaNet (henceforth EWN and BN). In the EWN (cf. Vossen 1998) additional semantic relations, i.e. Internal Language Relations, such as Role_Agent, Agent_Involved, Role_Instrument, Instrument_Involved, Role_Location, Location_Involved etc. have been introduced.

lish ten main relations holding between these roots and the corresponding suffixes, such as location (*prac-ov-iště* ‘workplace or study’), agentive-male (*prac-ov-nik* ‘worker’), agentive-female *prac-ovnice* ‘female worker’), adjective (*prac-ov-ni* ‘professional,’ ‘working’), derived nouns (*prac-ov-i-t-y* ‘diligence’) etc.

Although derivation is in many cases regular in the sense that it yields predictable results, certain irregular patterns show that it cannot be freely used for automatic semantic labelling and generation of derivational forms. Raffaelli et al. (2008) focus on derivational patterns of verbal aspectual pairs in Croatian and show that the semantic correspondence between imperfective and perfective verbs formed with the same affix (e.g. prefix *po-*) is far from regular (e.g. *jesti/pojesti* ‘to eat/to finish eating’; *skočiti/poskočiti* ‘to jump/to hop (once)’).

Hence, the elaboration of semantic relations between derivationally connected lexemes requires a closer inspection of each semantic field that is being processed (e.g. the Croatian suffix *-ač* is in some cases used for nouns denoting human agents, such as *pokretač* ‘instigator,’ *vozač* ‘driver,’ but also for nouns referring to instruments such as *pisač* ‘printer,’ *otvarač* ‘opener,’ etc.

Therefore, we believe that only through a thorough examination of individual semantic fields can one establish derivational patterns valid in different semantic fields, indicating specific semantic relations among morphologically related lexemes. We also believe that this approach will allow us to detect the particularities of individual semantic domains in terms of internal morphosemantic relations. However, the aforementioned examples show that the morphosemantic relations that could be widely applied in large computational lexicons such as WordNets require more elaborated theoretical foundations as well as practical analysis.

In this sense, we propose to apply a model of description in terms of morphosemantic fields such as those discussed and elaborated on by Raffaelli and Kerovec (2008). In our opinion, the morphosemantic field model offers a useful theoretical framework for the detection of morphosemantically related lexemes and the possibility of establishing semantic relations among derivationally connected lexemes in a further development of CroWN. To illustrate how the morphosemantic field model treats various semantic relations among morphologically connected lexemes, we will describe the morphosemantic field of lexemes referring to movement that are formed with the morphologically highly productive root lexeme **kre-*. We will discuss and graphically represent the semantic relations between the most frequent and salient lexemes in the morphosemantic field. This morphosemantic field has been chosen precisely because its root morpheme **kre-* is present in a large portion of the Croatian vocabulary denoting movement.²

2 According to the data gathered from the Croatian National Corpus: <http://www.hnk.ffzg.hr> and the Croatian National Repository: <http://www.riznica.ihjj.hr>.

2. Morphosemantic field model

The term morphosemantic field (henceforth MS field) was first introduced by P. Guiraud (1967). He defines an MS field “a set of lexemes integrated by a common etymon” and also as a “set of lexical units sharing the same lexical denominator” (1967: 125). The two definitions refer to two levels of linguistic description to which the MS field model is applied—the diachronic and the synchronic, respectively. From a diachronic perspective, Guiraud (1967) argues that the MS field model offers an etymological account based on the parallel (or complementary) development of the morphological and the semantic features of the derivatives. Synchronically, the MS field model can be used to examine relations between lexemes other than the solely paradigmatic ones described by the semantic field model (Raffaelli and Kerovec 2008: 143). According to Raffaelli and Kerovec (2008), lexemes that constitute an MS field share the same word base (morphological root or stem) and are members of different parts of speech (unlike the ones in a semantic field model and unlike the ones in a similar synset). On a conceptual level their meanings form an integrated cross-categoric structure gathered around a concept considered basic for their morphosemantic formation. The lexeme embodying this basic concept is referred to as the *basic lexeme* of an MS field (Raffaelli and Kerovec 2008).

Unlike the aforementioned authors (Koeva et al. 2008; Pala and Smrž 2004; Pala and Hlaváčková 2007), who mostly rely on formal approaches for the detection of morphosemantic relations, Raffaelli and Kerovec (2008) position the MS field model within the framework of cognitive linguistics stating that:

- a) MS fields are heterogenous, thus corresponding to the principles of prototype organization and lexical structures;
- b) they indicate a dynamic interplay of grammatical and semantic processes in structuring a vocabulary;
- c) cognitive processes metaphor and metonymy (Lakoff 1987; Lakoff and Johnson 1980; Langacker 1987) as well as generalization and specialization (Geeraerts 1997) are involved in the formation of new lexemes in an MS field.

Raffaelli and Kerovec (2008) also point out that metaphor/metonymy and generalization/specialization in an MS field motivate meanings across different parts of speech (e.g. a verbal meaning can motivate a nominal meaning) and not only within one word class (verbs for example). The present paper will also deal with the MS field in this way.

2.1. The concept of movement and the MS field *kretati se*

The concept of movement is represented in CroWN by different lexical hierarchies consisting of verbs, nouns and adjectives. Verbal lexemes denoting movement are divided into two main hierarchies. The first hierarchy contains the reflexive verbs *kretati se*, *gibati se*, *micati se/maknuti se* ‘to move oneself, to locomote,’³ denoting movement of various entities, whereas the second hierarchy contains the verbs *pokrenuti/pokretati*, *pokrenuti se/pokretati se*, *započeti/započinjati*, which can be interpreted as involving causes of movement. The semantic feature of causation is thus the main distinguishing feature of these two groups of verbs, e.g. *Vjetar se kreće kroz krošnje stabala* ‘The wind moves through the trees,’ where ‘wind’ is the object undergoing movement, as opposed to *Vjetar pokreće vjetrenjače* ‘The wind is moving the windmills,’ where ‘wind’ is the cause of movement. However, the meanings of *kretati se*, *pokrenuti/pokretati* ‘cause to move someone or something’ and the reflexive *pokretati se/pokrenuti se* ‘cause to move oneself, usually by conscious effort’ are morphosemantically closely related in the sense that the meaning of *pokrenuti/pokretati* and *pokretati se/pokrenuti se* is in fact motivated from the meaning of *kretati se* by the metonymic extension into the domain of causation (with the prefix *po-*). Such metonymic motivation is based on our knowledge of movement causes—all objects that undergo movement of some kind also presuppose the existence of an external or internal force that causes them to move (‘wind moves the windmills’, ‘gravity moves the celestial bodies’, ‘humans move themselves with their extremities, etc.). This highly salient metonymically motivated relation between *kretati se* and *pokrenuti/pokretati* cannot be in any way represented in CroWN since the verbs are clearly separated into different hierarchies. As this example shows, in order to describe and detect all the relevant semantic relations pertaining to the concept of movement, we must include the morphosemantic relations of various lexemes with the root lexeme *kre-*, although they might belong to rather different hierarchies in CroWN.

The basic lexeme of the MS field with the root lexeme *kre-* is the verb *kretati se* because it acts as a centre point around which other concepts are organized via metaphor/metonymy or specialization/generalization. Conceptually, this basic lexeme encompasses fundamental notions pertaining to movement of any kind—change of position through space by various means. This type of movement may be performed by all sorts of physical entities—humans, animals, vehicles, celestial bodies, fluids, etc. For example, *Asteroid se kretao ravno prema Zemlji* ‘An asteroid was moving straight towards the Earth’ vs. *Auto se kreće prevelikom brzinom* ‘The car is moving too fast’ vs. *Ribe se kreću pomoću svojih peraja* ‘Fish move with their fins,’ etc. Via schematization (Langacker 1987)

3 These and other definitions of the lexemes in this paper have been taken from the Croatian Language Portal: <http://hjp.srce.hr> and translated.

the verb itself acquires metaphorical meanings in more abstract spatially conceptualized domains; e.g. social (*Volio se kretati u umjetničkim krugovima* ‘He liked to move in artistic circles’) or scalar values (*Temperature se kreću od 20 do 30 stupnjeva za vrijeme ljeta* ‘Temperatures move from 20 to 30 degrees during the summer’). These semantic features of the verb *kretati se* provide foundations for the formation of new lexemes. These lexemes can belong to different parts of speech: nouns (e.g. *pokretnina/nekrenitna* ‘movable property/real estate’; *kretanje* ‘movement’), verbs (e.g. *pokrenuti/pokretati* ‘cause to move’) or adjectives (e.g. *pokretan* ‘movable’). The underlying processes motivating the formation of new lexemes and new conceptual structures are:

a. metonymic mappings, which highlight certain features within the same domain (in our case the spatial domain) and provide access to some conceptual properties by which the target concept is understood (Evans and Green, 2006), but also according to which the target concept is formed in an MS field in the first place. For example, *pokretnina* ‘movables, movable property,’ literally ‘that which can move’ (formed with the prefix *po-*), and its counterpart *nekretnina* ‘real estate’ (formed with the negative particle *ne-*), literally ‘that which does not move,’ are formed directly from the verb *kretati se* via metonymy. The metonymic motivation here occurs because, experientially, certain objects in one’s possession can undergo movement or can be moved from one location to another (cars for example), while others cannot (houses for example). Thus the ability to move as denoted by the verb *kretati se* becomes a salient and defining feature of one class of objects, ‘movables’ (denoted by the prefix *po-* plus the root *kre-*), while the constant lack of movement (denoted by the negative particle *ne-* plus the root *kre-*) is the defining feature of the other.

b. metaphorical mappings, i.e. cross-domain mappings (Lakoff, 1987; Evans and Green 2006), which are structured on the foundations of metonymic extensions of an MS field. For example, the verb *krenuti/kretati* ‘start moving, start going; go, head’ denotes the beginning of a movement leading to another location by a specific route or within a certain area. The metonymy and specialization processes motivate the formation of its meaning from the verb *kretati se* by limiting the whole action of movement to just one of its parts, namely the beginning of movement. This is shown in sentences such as *Krenuli su niz rijeku u svom čamcu* ‘They started going downstream in their boat,’ *Krenula je kroz šumu da dođe do bake* ‘She started going through the woods to get to her grandma,’ *Vlak za Split iz Zagreba kreće s perona 5 u 14 h* ‘The train for Split starts (i.e. leaves) from Zagreb at 2pm,’ etc. Next, from this metonymic basis of *krenuti/kretati* the salient features (the beginning of crossing a path) are abstracted and mapped onto other spatially conceptualized domains via metaphor. Hence, *krenuti/kretati* becomes synonymous with the verbs ‘to begin’ and ‘to start’ e.g. *krenuti večerati* ‘start eating dinner,’ *krenuti pjevati* ‘start singing,’ *krenuti pisati roman* ‘start writing a novel,’ etc. What this means is that actions

such as ‘singing,’ ‘writing’ or ‘thinking’ can be construed as sequenced processes having a beginning, duration and an ending. Their order of completion being similar to a path one must cross from the beginning to the end in order to obtain the desired result. Langacker (1987) refers to such construal of actions as goal-directed paths as a source-path-goal schema. In other words, parts of these actions are conceptualized as spatial points on a path, each one marking a certain degree of action completion. In this way the metonymic process in *krenuti/kretati* that maps the verb’s action onto the first part of a path (the beginning of movement) motivates the metaphoric extension which maps *krenuti/kretati* onto the first part of any action conceptualized as a path.

Another example of how metonymy and metaphor motivate new meanings is shown by the verbs *pokrenuti/pokretati* and *pokrenuti se/pokretati se*. As we have already mentioned, the causative meaning of the verbs *pokrenuti/pokretati* and *pokrenuti se/pokretati se* is motivated by an experience-based understanding of causal agencies initiating movement. This further motivates:

a. the meanings of a significant number of nominal forms via metonymy: the nouns *pokret* ‘a move, a gesture,’ e.g. *Jednim ga je pokretom srušio na pod* ‘With a single move he threw him on the floor,’ *pokretnost* ‘mobility,’ e.g. *Trbušni ples povećava pokretnost kukova* ‘Bellydancing increases the mobility of the hips,’ *nepokretnost* ‘immobility’ e.g. *Nepokretnost krajolika ga je sasvim očarala* ‘The immobility of the scenery enchanted him completely,’ and the adjectives *pokretan* ‘mobile’ and *nepokretan* ‘immobile’;

b. metaphoric extensions motivated by the causative features of the verbs *pokrenuti/pokretati* and *pokrenuti se/pokretati se* which metaphorically denote causation of any action and are thus synonymous with the causative verbs ‘to initiate,’ ‘to begin’ and ‘to start’ e.g. *Moj otac je pokrenuo vlastiti posao* ‘My father started his own business.’ The same metaphoric extensions motivate the meaning of the nouns *pokret* ‘an organized action of a social group’ (e.g. *komunistički pokret* ‘the communist movement’) and *pokretač* ‘initiator of an action’ e.g. *Televizijske kuće su česti pokretači humanitarnih kampanja* ‘TV companies are often initiators of humanitarian campaigns.’

Such metonymically and metaphorically motivated semantic relations are presented in the following diagram:

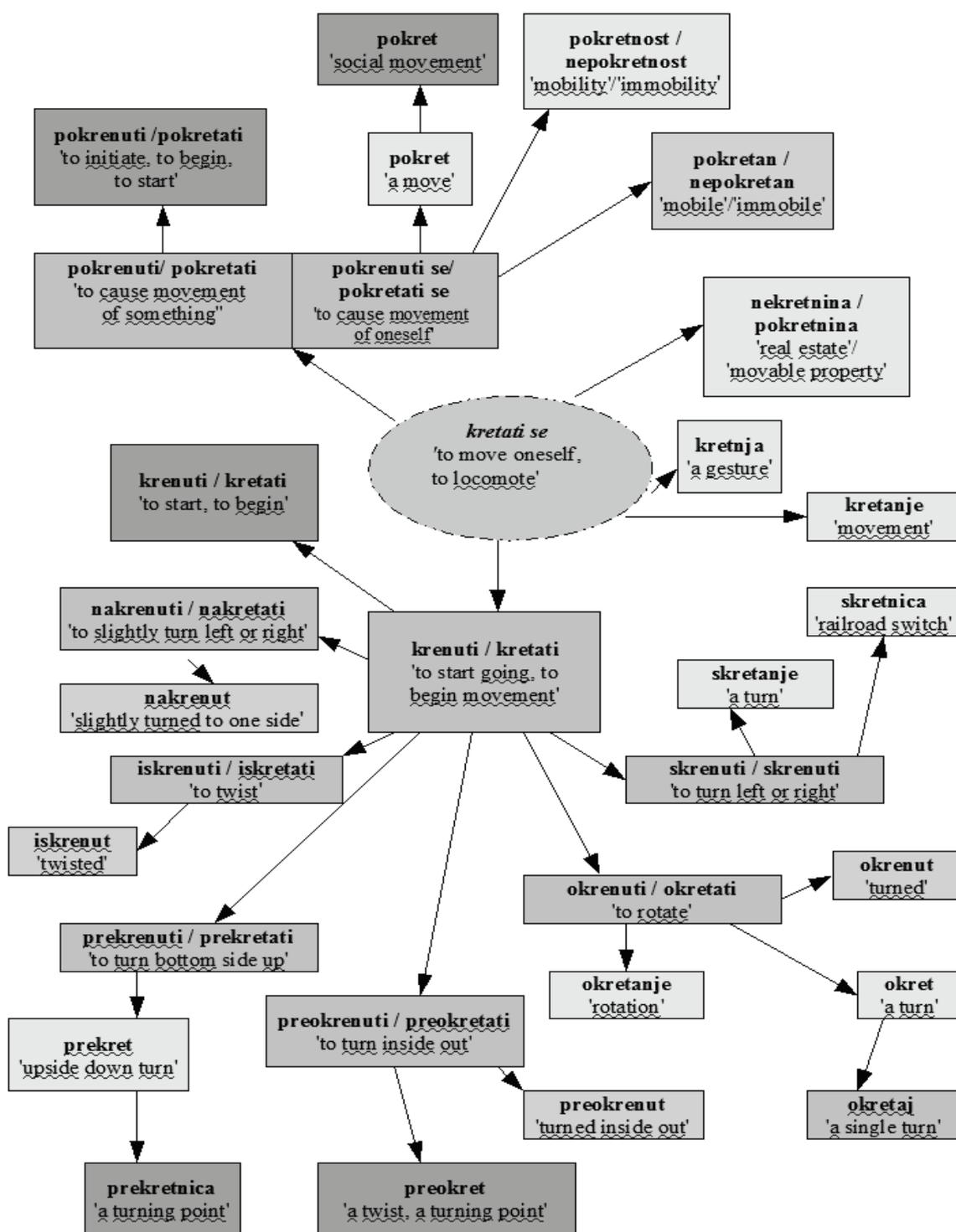


Diagram1. The MS field *kretati se*.

2.2. MS field *kretati se*

The above diagram shows how some of the most frequent derivatives are related to the basic lexeme and integrated into the MS field. We have already mentioned that metonymy and metaphor motivate the meaning of *krenuti/kretati* in terms of a goal-directed path and the corresponding source-path-goal schema. Furthermore, the metonymic and metaphoric processes motivate the formation of verb forms derived by prefixation (bottom centre part of the diagram) as well the formation of their corresponding nominal forms (left and right lower part of the diagram) from the verb *krenuti/kretati* by modifying the general features of the goal-directed path: every lexeme in this group denotes some type of deviation from a previously set course or position. For example, the verb *skrenuti/skretati* means ‘to turn; to make a turn, left or right’ (prefix *s-* plus the root *kre-*). It denotes a change of direction from a path previously established by the beginning of movement (*krenuti/kretati*) e.g. *Krenuo sam u centar grada, ali sam zabunom skrenuo u predgrađe* ‘I started off going to the centre of the town but I accidentally turned into a suburb.’ Because deviation from a previously set course is the salient feature of the meaning of *skrenuti/skretati* the metaphoric extension maps it onto other domains so that deviation becomes the salient feature of the verb’s meaning in other domains as well, usually perceived as a negative deviation from an accepted norm e.g. *skrenuti s uma* ‘turn away from sanity.’ Metonymy motivates the formation of the lexemes *skretanje* ‘a turn’ (e.g. ‘a turn on the road’) or ‘turning’ and *skretnica* ‘a shunt, a railroad switch’ as concepts experientially related to the action of turning left or right (a shunt, for example, controls the direction in which the train will head).

Such metonymic and metaphoric processes motivate the meanings of all these lexemes; each of them derived by a specific prefix activating different aspects of course/position deviation. We will describe two more examples of course/position deviation:

a. rotation around a vertical axis: the verb *okrenuti/okretati* ‘to turn around’ (e.g. *Okrenuo je glavu prema meni* ‘He turned his head towards me’) motivates the meaning of the noun *okret* ‘a turn’ (e.g. *Plesni okreti nisu uvijek jednostavni* ‘Dance turns are not always simple’) and the meaning of the gerund *okretanje* ‘rotation’ (e.g. *Okretanje Zemlje oko svoje osi razlog je zašto imamo dane i noći* ‘The turning of the Earth around its axis is the reason why we have days and nights’). Specialization of the noun *okret* then motivates the meaning of the noun *okretaj* ‘a turn; a single rotation’ in the sense that it limits the more abstract notion of ‘a turn’ to one of its instances—a single accomplished turn (e.g. *Koliko okretaja u minuti ima ovaj motor?* ‘How many times does this engine turn per minute?’). Also, metonymy motivates the meaning of *okretište* ‘the turning point for buses and trams’ in the same way the noun *skretanje* is motivated

by the verb *skrenuti/skretati*—*okretište* is the place where the action of turning is performed.

b. a change of position or of spatial properties of an object due to movement; denoted by verbs such as *nakrenuti/nakretati* ‘slightly turn diagonally left or right’ (prefix *na-*), e.g. *Nakrenula je glavu na lijevo* ‘She slightly turned (or slanted) her head to the left’); *iskrenuti/iskretati* ‘to twist’ (prefix *iz-*), e.g. *Iskrenula sam gležanj i sada me boli* ‘I twisted my ankle and now it hurts’ *prekrenuti/prekretati* ‘to turn something bottom side up’ (prefix *pre-*), e.g. *Prekrenula mi je ladice u stolu tako da su mi sve stvari poispadale* ‘She turned my desk drawers bottom side up so that all the things fell out of them’ and *preokrenuti/preokretati* ‘to turn something inside out’ (prefix *pre-*+ prefix *o-*) e.g. *Dobro je preokrenuti košulje prije peglanja* ‘It’s good to turn shirts inside out before ironing.’ The metonymic motivation of these meanings is based on the properties of objects that are affected by a specific type of movement—we know that a certain type of movement, such as *iskrenuti/iskretati* (twisting) has occurred because we see changes in the object’s position or properties before and after the movement (for example, the twisted ankle is now unnaturally turned to one side or swollen). This metonymy also motivates the meanings of the adjectives denoting properties of the object affected by motion: *nakrenut* ‘slightly turned; slanted,’ e.g. *nakrenuta slika* ‘a slanted painting,’ *iskrenut* ‘twisted,’ e.g. *iskrenut gležanj* ‘a twisted ankle’ and *preokrenut* ‘turned inside out,’ e.g. *preokrenut kaput* ‘a coat turned inside out.’ The verbs *prekrenuti/prekretati* and *preokrenuti/preokretati*, furthermore, motivate the meaning of the noun *prekret* and *preokret*, respectively, both nouns having a similar meaning of ‘an upside down turn’ (e.g. *Mrzim tjelesni jer nas tjeraju da radimo prekrete* ‘I hate gym class because they make us do upside down turns’). Because the verbs both denote a radical change of the spatial properties of an object due to movement (in a way that it changes the object’s basic properties—up is down and down is up/what was once on the inside has now been turned outside) the concept of radical change motivates metaphoric meanings of the nouns *prekretnica* ‘a turning point’ e.g. *Prvi svjetski rat velika je prekretnica u europskoj povijesti* ‘The First World War was a great turning point in European history’ and *preokret* ‘a twist; a turning point’ e.g. *Preokret u romanu nastaje ubojstvom glavnog lika* ‘The twist in the novel happens when they murder the main character.’ In both examples the nouns *prekretnica* and *preokret* denote a change that in some way alters the previous state of affairs which now are perceived differently, just as an object is perceived differently if it is for instance turned inside out (such as an item of clothing).

The description of the MS field *kretati se* illustrates the way in which morphosemantic processes work to build up the Croatian vocabulary. We believe that further theoretical and practical inquiries into the morphosemantic relations of various domains will show in other and detailed ways how the fundamental

conceptual and cognitive processes such as metonymy and metaphor, as well generalization and specialization, motivate new conceptual structures.

3. Conclusion and future work

The full integration of morphosemantic relations as presented above into the structure of a lexicon such as CroWN at the present moment seems like a challenging task in many ways. On the one hand, the theoretical framework of the MS field model can provide a far more extensive insight into the structure of the morphosemantic relations than the more formal approaches can. At the same time the MS field model, relying on the tenets of cognitive linguistics, is able to account for those semantic relations that can hardly, if at all, be explained by means of traditional morphological inquiries into the formation of meaning through derivation. In this way the complexity of morphosemantic relations becomes more explicit and its analysis more profound. On the other hand, the ways and extents to which the MS field model can be incorporated into a lexicological and lexicographic enterprise, such as the building of CroWN, remains to be seen, particularly in terms of its partially predefined structure. Nevertheless, we firmly believe that its full or partial application will prove to be useful in the course of our future work. We are also convinced that the model will be refined as a consequence of future research to be done in the area of morphosemantic relations within different semantic fields, thus allowing for more accurate formalization and establishment of regularities.

References:

- Evans, Vyvyan and Melanie Green (2006). *Cognitive Linguistics: An introduction*. Edinburgh: Edinburgh University Press.
- Fellbaum, Christiane, ed. (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.
- Geeraerts, Dirk (1997). *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*. Oxford: Clarendon Press.
- Guiraud, Pierre (1967). *Structures étymologiques du lexique français*. Paris: Larousse.
- Koeva, Svetla, Cvetana Krstev, and Duško Vitas (2008). Morpho-semantic relations in WordNet – a case study for two Slavic Languages. In *Proceedings of the Fourth Global WordNet Conference*, edited by Attila Tanács, Dóra Csendes, Veronika Vincze, Christiane Fellbaum, and Piek Vossen, 239–253. Szeged: Global WordNet Association.
- Lakoff, George (1987). *Women, Fire and Dangerous Things. What Categories Reveal About the Mind*. Chicago: Chicago University Press
- Lakoff, George and Mark Johnson (1980). *Metaphors We Live by*. Chicago – London: University of Chicago Press.

- Langacker, Ronald W. (1987). *Foundations of Cognitive Grammar. Volume 1*. Stanford: Stanford University Press.
- Pala, Karel, and Pavel Smrž (2004). Building Czech WordNet. *Romanian Journal of Information, Science and Technology* 7:1–2: 79–88.
- Pala, Karel, and Dana Hlaváčková (2007). Derivational relations in Czech WordNet. *Proceedings of the Workshop on Balto-Slavonic natural Language Processing*, 75–81. Prague: ACL.
- Raffaelli, Ida, Marko Tadić, Božo Bekavac, and Željko Agić (2008). Building the Croatian WordNet. In *Proceedings of the Fourth Global WordNet Conference*, edited by Attila Tanács, Dóra Csendes, Veronika Vincze, Christiane Fellbaum, and Piek Vossen, 349–359. Szeged: Global WordNet Association.
- Raffaelli, Ida and Barbara Kerovec (2008). Morphosemantic fields in the analysis of Croatian vocabulary. *Jezikoslovlje* 9.1–2: 141–169.
- Vossen, Piek (1998). *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Dordrecht: Kluwer.

Sources:

- Croatian National Corpus: <http://www.hnk.ffzg.hr>.
Croatian National Repository: <http://www.riznica.ihjj.hr>.
Croatian Language Portal: <http://hjp.srce.hr>.