UDK 811.163.42'374.2
81'374.2:165.194
Izvorni znanstveni članak
Prihvaćeno za tisak: 4. studenoga 2015.

Bruno Nahod
Institute of Croatian Language and Linguistics, Zagreb
bnahod@ihjj.hr

Domain–specific Cognitive Models in a Multi–Domain Term Base

More recent approaches to terminology incorporate premises from cognitive linguistics and sociolinguistics in their theoretical framework. Although such approaches range from general theoretical frameworks to more practical models based on a specific cognitive theory, none provides explicit guidelines on how to incorporate theoretical findings into a multi–domain term base, such as Struna. Struna (Stojanov et al. 2009; Struna 2012) is a national database of Croatian Special Field Terminology. Since the Struna database includes a wide array of specialized domains, this entails the design of a new sociocognitive model[1] that integrates premises from modern terminology theory and cognitive sociolinguistics into a usable practical framework for a multi–domain term base.

1. Introduction

Sociocognitive approaches to Terminology are not a novelty. These theories, which arose in reaction to the General Theory of Terminology (GTT) (Wüster 1979), incorporate premises from cognitive linguistics and sociolinguistics (Faber 2009: 110). However, none provides guidelines that explain how to incorporate theoretical findings into a multi–domain term base (Nahod and Vukša 2014).

Struna is the Croatian National Termbank (Struna 2012), which will eventually include terminological units belonging to most specialized knowledge domains in Croatian language. Until now, the processing of terminological units has been generally performed based on the principles of Wuster's General Terminology Theory (GTT)[2]. Nevertheless, as more domains have been added, certain problems have emerged.

1 Research reported in this paper was co–financed (286,505.32 HRK) by the European Social Fund within OP *Human Resources Development* 2007–2013.

2 Since the topic of this paper requires a multi–faceted theoretical and practical framework, we will have to take the limitations of space into consideration; therefore, we will not be able to get into the details on how and why Struna was established. For papers on this topic, please refer to Bratanić and Ostroški (2013) as well as Bergovec and Runjaić (2012). For the same reason we will not be going into the problems of the GTT, especially considering the ample literature on this subject.

Eighteen domains have now been processed and made accessible to the general public with an additional four in various stages of processing. The biggest problem that we are currently facing is that of conceptual variation. More specifically, there are terminological units that belong to and are used in two or more specialized languages. Thus, the same term designates slightly or even radically different concepts as reflected in their definitions and relations with other concepts.

A great deal of effort and energy was invested in trying to find a solution for this harmonization problem (Bergovec and Runjaić 2012). For this purpose, we followed the recommendations for term/concept harmonization in ISO 860_2007 and 704:2009 but with little success. Among the examples reported by Bergovec and Runjaić is the case of **anode**. Even after harmonization, its definition still varies, depending on the specialized knowledge domain (i.e. physics, engineering, chemistry, or chemical engineering).

domain	definition
Physics	electrode with greater electrical potential
Engineering	positive electrode in an electrolytic cell
Chemistry	negative–charge electrode
Chemical Engineering	electrode with predominately anode reaction

Table 1 Definitions of anode in Struna

These recommendations for harmonization were unable to solve cases such as **anode** because they stem from the GTT, which is based on the classical theory of categorization. As highlighted by Lakoff, the classical theory of categorization assumes that human categories are more or less automatized derivations of objective categories that exist in the real world (Lakoff 1987: xii). *Automatized derivation* signifies that category organization is based on characteristics that exist in a real world. It is assumed that entities belong to the same category if and only if they share certain properties. These properties are the necessary and sufficient conditions for defining the category (Murphy 2002: 11). The influence of the classical theory of categorization on the GTT can be observed in Felber's description of intrinsic and extrinsic characteristics:

> Intrinsic characteristics are more convenient than extrinsic characteristics because they can be ascertained by inspection only because they are self–sufficient (Felber 1984: 119).

Rosch[3] (1975; 1978) challenged this classical view, which led to new proposals that were more focused on empirically confirmed theories of human cognition. Despite certain differences, subsequent theories of categorization

3 Prior to Rosch's pioneering work, other scholars also argued against objectivism i.e. Hanson (1961), Hesse (1963), Kuhn (1970) and Wittgenstein (1953).

agree that categorization is based on human perception and consequently much less objective than it was assumed.

2. (Neuro) cognitive studies of specialized language

Based on our analysis of terminological units in the Struna database, we hypothesized that expert users of a specialized language (experts in a knowledge field or motivated domain[4]) and non–expert users would process terminological units differently. It is our assertion that the mechanisms that facilitate knowledge acquisition in a motivated domain are different than those used in everyday life. Categorization and the acquisition of general knowledge depend on human interaction with the physical and sociocultural environment. However, the categorization that underlies specialized language acquisition and use is often the product of conscious and motivated learning, and thus a more goal–directed.

Consequently, conceptual structures in the minds of experts in terms of broadness, complexity and the stability of the substructures could be different from those in the minds of non–experts. Lakoff (1987: 334) states: “For me, conceptual systems with different organizations are different systems”. If a significant difference in conceptual organization between experts and non–experts indeed exists, it would be necessary to conceive a paradigm specifically for describing conceptual structures in specialized language.

In this sense, of special interest are expert–novice (laymen) studies that focus on differences in categorization and conceptual processing in user groups with different levels of expertise.

Experts vs. non–experts (novices)

There are various studies on the difference in categorization between experts and non–experts. Unfortunately, most compare “lab–made” experts and non–experts. Such experiments are designed in such a way that the ’expert group’ undergoes training in which they learn abstract categories such as geometrical shapes, colors, and graphical patterns (Hoffman, Harris, and Murphy 2008; Kaplan and Murphy 2000; Lin and Murphy 1997; Palmeri and Blalock 2000; Palmeri and Blalock 1999 *inter alia*), and say little about aspects related to the type of lifelong learning acquired by a true expert.

Murphy and Wright (1984) reported the results of experiment conducted on four groups of subjects with different levels of expertise[5]. The experiment consisted of listing the typical characteristics of an aggressive child, a depressive child, and a disorganized child. The authors reported statistically relevant differences in the number of characteristics listed and in the level of the agreement within a group. Both variables corresponded to the level of expertise.

4 A motivated domain is any knowledge domain which a person is motivated to learn because of higher education, professional specialization, personal hobbies or interests.

5 In clinical psychology.

Not surprisingly, experts had the highest number of characteristics listed, and the highest level of agreement within the group was observed, while novices had the lowest results for both variables (1984: 147).

Shafto and Coley (2003) studied the nature of categorization strategies in experts and novices[6]. They examined the degree to which taxonomic similarity and domain–specific knowledge influenced how experts and novices categorized and reasoned about concepts that could be perceived both as domain specific and as general. Shafto and Coley observed that experts were sensitive to the property being generalized in a way that non–experts were not. Their results indicated that experts have access to domain–specific conceptual relations that are unknown to non–experts. Experts do not use these relations as the sole means for reasoning, but instead rely on them as an alternative to taxonomic similarity only when they are considered relevant (2003: 647). Their conclusion was that expertise involves the capacity for perceiving deeper reasons underlying this similarity as well as for learning new domain–specific relations that are potentially orthogonal to taxonomic relations, and knowing when to rely on one type of relation instead of another (2003: 648).

Explicit and implicit learning

If there is a difference in categorization between experts and non–experts, then the question arises of whether the brain supports this kind of parallel cognitive function. Studies of patients with category–specific deficits clearly show that parallel systems associated with categorization do exist. These include living things compared to nonliving things (Mahon and Caramazza 2009; Capitani et al. 2003), fruits compared to vegetables and/or animals (Caramazza and Shelton 1998), various subsystems for the visual properties of living things (Moss, Tyler, and Jennings 1997), and different content and structure of concepts in various semantic categories (Taylor et al. 2001).

Furthermore, clinical lesions in the medial temporal lobe (MTL) in monkeys corresponding to lesions in patients have reproduced many properties of human impairments (Mahut and Moss 1984; Mishkin 1982; Mishkin 1978; Squire and Zola–Morgan 1983). Both in monkeys and in human patients, declarative memory impairments were observed, while functions that control skill acquisition and other functions associated with non–declarative memory remained intact.

Reber et al. (2003) reported evidence for separate learning systems in two groups of healthy subjects, depending on whether patients were learning consciously. Their study showed that both groups of subjects managed to learn new categories although only one was informed that the stimuli were organized in categories. Furthermore, their fMRI data showed that implicit and explicit category learning activates different brain areas.[7]

6 In marine fauna.

7 When the category was learned implicitly, decreased occipital activity was observed for categorical patterns. Explicit acquisition of the category, on the other hand, showed increased activity for categorical patterns in the left occipito–temporal cortex, right anterior prefrontal cortex, and medial parietal cortex (Reber et al. 2003: 578).

A recent fMRI study on the neurolinguistic substrate of specialized knowledge processing by Faber et al. (2014) investigated the difference between experts[8] and non–experts in a task involving matching related words to target stimuli. This experiment produced evidence that the system for expert knowledge seems to be directly related to the activation of supramodal representations recovered from the temporal area as well as to visual scene generation encoding meaningful contexts (Faber et al. 2014: 19).

Such studies provide support for the assumption that categorization in specialized language could be based on different mechanisms than categorization in general language[9]. This would signify that conceptual structures in experts are different, not only in the number of concepts but also in the relations between concepts.

Sociocognitive terminology today

As mentioned in the introduction, sociocognitive approaches to Terminology are not a novelty. Most of them have a solid theoretical background in modern paradigms of human cognition and/or semantics. Although the purpose for developing a new approach was not to contradict or criticize other cognitive approaches, the specific needs and problems that emerged from terminological work in Struna (see Bratanić and Lončar 2015; Bergovec and Runjaić 2015) resulted in a certain divergence both in practical and theoretical aspects. Frame–based terminology (León Araúz, Faber, and Montero Martínez 2012; Faber, Márquez Linares, and Vega Expósito 2005; Faber et al. 2006) is a highly developed framework for structuring specialized domains and creating non–language–specific representations (Faber 2009: 121) that is based on Fillmore's Frames (Fillmore 1985). Even though one cannot question the quality and usability of the proposed framework, we have found that implementation of such detailed designed templates into our database exceeds our resources, as well as the needs of our users. However we believe that since they approach the conceptual structures on different levels (frame–base on the intradisciplinary level and our sociocognitive model on the interdisciplinary level), it is plausible that certain level of compatibility could be achieved.

Despite the similarity in name and in certain basic premises, comparison of the Sociocognitive Theory of Terminology (STT) proposed by R. Temmerman (2000) and our sociocognitive model reveals many practical and theoretical differences. Temmerman investigates conceptual variations and category structure in the "special language of life sciences" (2000: 46). As the result of her investigation, Temmerman concludes that most of the categories have a prototypical structure and that the way the concept is described may vary depending on numerous parameters, a representation of which can be explained in the form of cognitive models.

8 In geology.
9 Preliminary results of our own study (yet to be published as a full report) strongly support these findings (Nahod and Vukša 2014; Nahod 2015).

Although our approach is based on certain aspects of the STT, the obvious one being the use of cognitive models, our initial premises diverge from those of the STT, and this resulted in an evident difference in paradigms. Temmerman uses the premises of cognitive linguistics that emerged from research on general language. In contrast, our research is based on the hypothesis that categorization and conceptualization processes in the Language for Special Purposes are somewhat different from the ones in general language. This signifies that the premises of sociocognitive linguistics should be attuned accordingly.

In our opinion, the usage of the cognitive models requires a more pronounced social component in a sense that a homogeneous community should be viewed as a determiner for their functionality. In this view the field of life sciences, since it includes numerous sub disciplines, cannot be seen as a domain with only one specialized language, but rather as a collection of multiple domains of closely related yet different specialized languages. We believe that the reported prototype structure of the categories, a phenomenon not observed in our study, is the result of leveling the different specialized languages of life sciences, which induced the emergence of general language properties in the analysis. We have classified this aspect of LSP as a general language component, and have tried to describe it as a crucial but segregated aspect of specialized language description by introducing a general language module.

3. Towards a new terminographic model of describing conceptual structures

As previously mentioned, the configuration of terminological units in Struna is mostly based on GTT terminographic principles. The inclusion of new domains over the last six years has created an increasing number of problems. Apart from multiple entries, there is also the fact that terms in the definitions of one domain do not necessarily correspond to the same concepts in definitions in other domains.

For example, in domain D_1, there is a term T_1 that designates concept C_1, whose definition is in the form of *genus proximum et differentia specifica*. Quite understandably, the characteristics in this definition are domain–specific. These characteristics reflect the collective knowledge of experts in that domain and, thus, the conceptual structure underlying that knowledge.

However, in a second domain D_2, term T_1 is also used to designate the characteristics of various concepts. Unfortunately, term T_1, when used in D_2, may not designate the same concept as in D_1. In fact, it may not even be used as a term and be simply a lexeme taken from the general language lexicon. Although in theory, all of the terms used in definitions should be directly related to the designated concept, we have found that following this rule closely would exponentially increase the problem of multiple entries. In our opinion, this kind of inter–domain conceptual variation should be addressed in term bases. Accordingly, it should be described in a coherent way in order to avoid the confusion stemming from multiple results that seem to be of equal value.

The solution lies in the design of a new model for terminographic processing. This proposal is based on Idealized Cognitive Models (ICM) (Lakoff 1987). Our research showed that experts from different fields of knowledge often perceive the same concept in a slightly different or completely different manner. The case of **anode** is just one of many.

These differences, which are perceived as conceptual variation if they are at the term base level, reflect domain–specific categorization, and are relatively stable for the group of experts with the same knowledge domain. Furthermore, the definition of a concept can often entail different definitions of other concepts in the same domain–specific categorization.

Experts in a specialized field can be regarded as subgroup of society or a subcultural community that transcends definitional boundaries such as language, culture, and personal beliefs or preferences. This kind of conceptual variation in the specialized language used by a group can be better understood if it is described using cognitive models. For this purpose, we propose Domain–specific Cognitive Models (DCMs) as schemata for inter–spheres in which certain specialized concepts form relatively stable relations and are categorized in different degrees of specialization in regards to their counterparts in other DCMs.

DCM can be perceived as a collective conceptual subsystem consisting of concepts that show variation in properties that define them in comparison to general or more broadly accepted conceptual systems. This variance, emerges either as a result of highly specialized (deeper) knowledge or because the research is focused on specific properties (non–intrinsic) of the subject.

Examples and implementation

Anode/cathode (hr. *anoda/katoda*) and **space** (hr. *prostor*) are concepts to which DCMs can be applied. Both **anode**[10] and **cathode** can be understood and categorized differently, depending on the knowledge domain. In general language, **anode** is a positive electrode, and **cathode** is a negative electrode. Although this is also true in most specialized domains, there are exceptions. The analysis of these terminological units reveals three defining characteristics of **electrodes**:

1.	electrical charge	positive or negative
2.	type of reaction	anodic or cathodic
3.	electric potential	greater or smaller

Table 2 Three approaches to defining an electrode

These characteristics are not always mutually exclusive because **anode** can have a positive electrical charge' (1), an 'anodic reaction' (2), and a 'greater electric potential' (3). If these were the only characteristics, then the term

10 Since anode and cathode are always opposites, we will use only anode in our examples.

anode would always designate the same concept, though defined in a different way – using a different set of characteristics. Nonetheless, a greater electric potential does not necessarily imply a positive charge. Therefore, one of the **anodes** could be defined as the antonym of the other two. As previously mentioned, in chemistry, **anode** is defined as a negative electrode. Consequently, the definitions of related concepts such as **anodic reaction** depend on the definition of the anode. If the anode is negative, then it will attract positively charged particles, and if it is positive, it will attract negatively charged particles[11]. Not surprisingly, it is impossible to describe this kind of complex conceptual relationships in a term base based on GTT principles.

As a way of circumventing this harmonization problem (see Bergovec and Runjaić 2012), Struna now includes two parallel "branches" for the categorization of **anode**:

Positive: for electrical engineering and general physics

Negative: for chemistry and radiology (physics).

Moreover, *greater electrical potential* and *anodic reaction* belong to either (or both) categorizations. In the Struna database, there are 182 terminological units that are in some way related to **anode**, five of which have explicitly defined IS–A relations. Most of the other relations, however, are impossible to specify.

Nevertheless, all of the inter–conceptual relations could be described if a DCM were superimposed onto the existing conceptual structure. The use of four correlated Domain Cognitive Models (DCMs) would make it possible to describe all of the cases of **anode**, independently of their knowledge domain. More importantly, it would be possible to optimally define all of the relations between existing terminological units related to **anode**, and also provide a stable definitional structure for future ones.

A users search for "anode" in Struna would than generate all the layers of categorization in which concept is found. By describing conceptual structures in terms of DCMs, concepts can be "grouped" in layers that better represent inter–field and intra–field conceptual relations. In the case of **anode**, this would require three DCMs[12].

It goes without saying these layers would not be limited to individual concepts, such as **anode** or **cathode**, but would include all the concepts that particular DCM invokes. For example, a layer of "positive anode" would include all concepts that have stable relations within the DCM such as the following: **cathode** (negative one), **anodic reaction, cathodic reaction, electrolysis**, etc.

The implementation of this kind of description in our term base is far from simple. There are at least two issues which need to be addressed prior to this. First of all, top and bottom layers must be defined. Evidently, there

11 This explanation is simplified for the purposes of this paper.
12 The number of DCM layers should be flexible.

are many ways of classifying such concepts, namely, by theories, by branches of the field, by schools of thought, by subjects of study, etc. Furthermore, too many layers could result in a representation so complex that it would not be helpful or useful to the end user. Precisely for those reasons, it would be a mistake to arbitrarily choose a given number of layers since no two domains have conceptual systems of the same size or complexity.

For example, the field of physics cannot be compared with that of corrosion and protection of materials[13]. An in–depth analysis of the specialized knowledge units would have to be performed prior to the implementation of layers. The definition of the top layer thus requires the implementation of the general language module since it is evident that the specialized language stems from general language. Therefore, a general language module (GLM) should be the logical choice for top–level layer. Similar solutions can be observed in the ’flexible definitions’ introduced in EcoLexicon (San Martín and León–Araúz 2012), and possibly in Temmerman’s description of a ’unit of understanding’ (2000: 224).

The second issue is that the identification of layers or Domain Cognitive Models, which requires a knowledge of the field that generally surpasses that of the average terminographer. Therefore, close collaboration with field experts will have to be included in this phase.

A sketch of implementation

This section provides examples of possible solutions from an early outline of the formally described implementation of the DCM in a multi–domain term base (pre–alpha). The schema in the examples is simplified and only shows the most basic elements.

```
<?xml version="1.0" encoding="UTF-8"?>
<martif>
<martifHeader>
    <fileDesc version="pre-alpha">
        <titleStmt>
            <title>Struna - Domane Cognitive Model - sample</title>
        </titleStmt>
</martifHeader>
<text>
    <body>
      <GLM altName="General Language Module">
      <domain id="01" name="fizika" altName="physics">
      <domain id="02" name="antropologija" altName="anthropology">
      <domain id="03" name="kemija" altName="chemistry">
```

Figure 1. Top–level elements

13 “Corrosion and Protection of Materials” was a terminological project in Struna that the author worked on in 2009 and 2010, which resulted in 892 terminological units that were included in the term base. In comparison, the project of the “Terminology of physics” produced 3507 terminological units.

Figure 1 shows the top–level elements in the structure of the term base. Along with the specialized knowledge domains formalized by the element <domain>, we have implemented the General Language Module <GLM> which will serve as a neutral top–level for our cognitive models.

```
<GLM altName="General Language Module">
    <metData>
        <source>
        <layers>
            <layer id="1_00" extendedness="neutral"/>
            <layer id="2_00" extendedness="figurative"/>
        </layers>
    </metData>
    <lexicon>
        <lexCat id="0001">
            <lex>prostor</lex> <!-- EN "space" -->
            <layer id="1_00">
                <sense id="1" type="general">
                <sense id="2" type="exact" context="površina">
                <sense id="3" type="exact" context="površina">
            </layer>
            <layer id="2_00" type="metaphor"/>
        </lexCat>
        ...
    </lexicon>
</GLM>
```

Figure 2. Structure of the GLM

The main function of the GLM is to include lexemes that designate broader concepts in definitions. The General Language module has a literal layer and a figurative layer (Figure 2). The figurative layer includes the metaphoric senses of the concept (Figure 4).

Specialized domains are also structured into layers. Unlike the GLM, the number of layers in domain structure is not defined but reflects the number of recognized DCMs in the domain.

```
<domain id="01" name="fizika" altName="physics">
    <metData>
        <layers>
            <layer id="1_01" name="DCM_PF1" extendedness ="classical"/>
            <layer id="11_01" name="DCM_PF11" extendedness ="radiology"/>
            <layer id="2_01" name="DCM_PF2" extendedness ="modern"/>
            <layer id="21_01" name="DCM_PF21" extendedness ="relativistic"/>
            <layer id="22_01" name="DCM_PF22" extendedness ="quantum"/>
            ...
        </layers>
    </metData>
    <CS altName="Conceptual Structure">
        <termUnit id="0001" layer="1 01">
        <termUnit id="0002" layer="2 01">
        <termUnit id="0003" layer="1 01">
        <termUnit id="0004" layer="21 01">
        <termUnit id="3210" layer="21 01">
        <termUnit id="3211" layer="21 01">
        <termUnit id="3212" layer="21 01">
        <termUnit id="3213" layer="21 01">
        <termUnit id="9801" layer="1 01">
        <termUnit id="1269" layer="11 01">
    </CS>
  </domain>
```

Figure 3. GLM layers in the domain of Physics

An analysis of terminological units related to **anode** /**cathode** and **space** revealed that the conceptual system in Physics can be structured into two high–level DCMs[14]: classical (id=”1_01”) and modern (id=”2_01”). Furthermore, two sublevels of modern Physics are relativistic (id=”21_01”) and quantum physics (id=”22_01), whereas classical physics includes the level of, radiology (id=”11_01”).

As shown in Figure 3, the specification of the corresponding DCM for each terminological unit using argument “layer” enables us to group concepts into semantic layers, a feature that is not easily implemented in term bases whose organization is based on GTT principles.

14 We are basing this approximation on the terminological units in Struna. Physics may (and will) eventually include more than four DCMs.

```
<GLM altName="General Language Module">
    <metData>
        <source>
                <name>HJP</name>
                <url>http://hjp.novi-liber.hr/</url>
            </source>
        <layers>
            <layer id="1_00" extendedness="neutral"/>
            <layer id="2_00" extendedness="figurative"/>
        </layers>
    </metData>
    <lexicon>
        <lexCat id="0001">
            <lex>prostor</lex> <!-- EN "space" -->
            <layer id="1_00">
                <sense id="1" type="general">
                    <definition>neograničena protežitost, ukupnost odnosa u svim
                            dimenzijama i pravcima</definition>
                </sense>
                <sense id="2" type="exact" context="površina">
                    <definition>dio Zemljine površne različite veličine</definition>
                </sense>
                <sense id="3" type="exact" context="površina">
                    <definition>ograničen dio površine na koji se može što
                            smjestiti</definition>
                </sense>
            </layer>
            <layer id="2_00" type="metaphor"/>
        </lexCat>
        ...
    </lexicon>
</GLM>
```

Figure 4. Lexeme space in GLM[15]

Since the concept of **space** is one of the most complex in Struna, it was used as a starting point for the development of our model. Figure 2 shows the lexeme 'space' schematized in the GLM. The description includes three typical senses of 'space' in the Croatian language, plus a figurative layer for more metaphorical uses[16].

The system of identification codes (ids) makes it possible to access the meaning of the lexeme at any level. Figure 4 depicts the relation between the definition of **space** in classical physics and its definition in general language: <target interLink="001" id="0001" lex="prostor" sense="1"/>, where the 'interLink' value refers to layer 1 of the GLM; the 'id' value refers to the lexeme id; and the 'sense' value to the sense id.

15 Translations: 1. in "general" unlimited expanse, the totality of relations in all dimensions and directions; 2. in "exact" id="02" part of the Earth's surface of different sizes; 3. in "exact" id= "03" a limited part of the surface on which something can be placed.

16 The elements of the GLM figurative layer have not as yet been formalized.

```xml
<CS altName="Conceptual Structure">
    <termUnit id="0001" layer="1_01">
        <langSet xml:lang='hr-HR'>
            <term type="preferredTerm">prostor</term>
            <term type="fullForm" >trodimenzijski prostor</term>
            <definition>neograničena trodimenzijska protežnost u kojoj tijela
                        imaju relativne položaje</definition>
            <relations>
                <relation id="01" type="simmilar to">
                    <target interLink="1_00" id="0001" lex="prostor" sense="1"/>
                </relation>
                <relation id="02" type="broader">
                    <terget interLink="2_01" id="0002" term="n-dimenzijski prostor"/>
                </relation>
                <relation id="03" type="hypernym">
                    <target interLink="1_01" id="0003" term="euklidijski prostor"/>
                </relation>
            </relations>
        </langSet>
        <langSet xml:lang='en-EN'>
            <term>space</term>
        </langSet>
        </termUnit>
```

Figure 5. **Space** in the ’classical physics’ layer[17]

Argument type=”similar to” in the ’relation’ element refers to the type of the relation, and indicates that the definition of **space** in classical physics corresponds to the general meaning of **space** in general language. The same syntax with its corresponding arguments was used to define the relations between the definitions of **state of exception** (stanje izuzeća) and **workshop** (radionica) from the domain of Anthropology (Figure 7).

17 Translations: 1. “PreferredTerm” space; 2. “fullForm” three–dimensional space; “definition” unlimited expanse in which bodies have relative positions; 3. “lex” space; 4. “term” id=”0002” n–dimensional space; “term” id=”0003” Euclidian space.

```
<domain id="02" name="antropologija" altName="anthropology">
    <metData>
        <layers>
    </metData>
    <CS>
        <termUnit id="0802" layer="1_02">
            <langSet xml:lang='hr-HR'>
                <term type="preferredTerm">stanje izuzeća</term>
                <definition>
                    <ref type="GLM_fgr">
                    <target interLink="2_00" id="0001" lex="prostor">prostor</target>
                    </ref> izvan političke zajednice u koji suverena vlast isključuje
                    izvan zakona one svedene na goli život</definition>
            </langSet>
            <langSet xml:lang='en-EN'>
                <term>state of exception</term>
            </langSet>
        </termUnit>
        <termUnit id="0865" layer="3_02">
            <langSet xml:lang='hr-HR'>
                <term type="preferredTerm">radionica</term>
                <definition>
                    <ref type="GLM_ntr">
                    <target interLink="1_00" id="0001" lex="prostor" sense="3">
                    prostor</target>
                    </ref> unutar kojega specijalizirana radna snaga proizvodi ...
                    predmete namijenjene razmjeni</definition>
            </langSet>
            <langSet xml:lang='en-EN'>
                <term>workshop</term>
            </langSet>
        </termUnit>
```

Figure 6. 'State of exception' and 'workshop' (Anthropology)[18]

In the case of **workshop**, <target interLink="001" id="0001" lex="prostor" sense="3"> refers to the third sense of **space** in general language, with the meaning "limited part of the surface on which you can place something". As mentioned earlier, 'space' in the definition of **state of exception** is actually a metaphor. Accordingly, the relation <target interLink="002" id="0001" lex="prostor"> directs us to the figurative layer of **space**.

Apart from cross–domain relations, the syntax permits relations between different layers of the same domain. Figures 13 and 15 show various types of conceptual relations in different DCMs in physics (broader, hypernym, narrower, hyponym, etc.), for which the value of the argument 'interLink' determines the specific layer of the domain. For example, 'interLink="**01**21"' in the second relation for the **n–dimensional space** refers to the domain 'physics' (0121), sublayer 'relativistic' in modern physics (0121).

The addition of the argument 'usgLayer' is a good solution for the previously described problem of **space** (Nahod 2015: 175). In the <term type="short form" usgLayer="012">prostor</term> for the **n–dimensional space**, the value of 'usgLayer' indicates that the short form **space** (*prostor*) is used for this concept in the DCM that we have called 'relativistic' (Figure 7).

18 Translations: 1. id="0802" "definition" space outside the political community in which sovereign power excludes from law those that have been reduced to bare life; 2. id="0865" "definition" space in which specialized labor force produces artifacts for the trade.

```
<domain id="01" name="fizika" altName="physics">
    <metData>
    <CS altName="Conceptual Structure">
        <termUnit id="0001" layer="1 01">
        <termUnit id="0002" layer="2_01">
            <langSet xml:lang='hr-HR'>
                <term type="preferredTerm">n-dimenzijski prostor</term>
                <term type="short form" usgLayer="2_01">prostor</term>
                <definition>neograničena n-dimenzijska protežnost u kojoj tijela
                        imaju relativne položje</definition>
                <relations>
                    <relation id="01" type="narrower">
                        <terget interLink="1_01" id="0001" term="prostor"/>
                    </relation>
                    <relation id="02" type="hyponym">
                        <target interLink="21_01" id="0004" term="prostorvrijeme"/>
                    </relation>
                </relations>
            </langSet>
            <langSet xml:lang='en-EN'>
                <term>n-dimensional space</term>
            </langSet>
        </termUnit>
```

Figure 7. N–dimensional space[19]

By implementing cross–domain relations, we were able to explicitly connect terms that appear in definitions of one domain to the terminological units in other domains. Situations that require such solutions tend to appear in multi–field domains such as physics, chemistry, medicine etc.

For instance, the concept **anode** is defined in physics as a positive electrode, a definition that corresponds to most of the other concepts connected to **anode** in the same domain.

19 Translation: “definition” infinite n–dimensional extent in which bodies have relative positions.

```
<domain id="03" name="kemija" altName="chemistry">
    <metData>
        <layers>
            <layer id="1_03" name="DCM_CH1" extendedness ="general"/>
            <layer id="2_03" name="DCM_CH2" extendedness ="chemical engineering"/>
            <layer id="21_03" name="DCM_CH21" extendedness ="galvanic cells"/>
                    ...
        </layers>
    </metData>
    <CS>
        <termUnit id="3670" layer="21_03">
            <langSet xml:lang='hr-HR'>
                <term type="preferredTerm">anoda</term>
                <definition>negativno nabijena elektroda</definition>
                <relations>
                    <relation id="01" type="altCat">
                        <terget interLink="1_04" id="5437" term="anoda"/>
                    </relation>
                    <relation id="02" type="altCat">
                        <terget interLink="1_05" id="1237" term="anoda"/>
                    </relation>
                    <relation id="02" type="diffCat">
                        <target interLink="11_01" id="9801" term="anoda"/>
                    </relation>
                </relations>
            </langSet>
            <langSet xml:lang='en-EN'>
                <term>anode</term>
            </langSet>
```


Figure 8. **Anode** in Chemistry[20]

There are certain subfields in Physics that correspond to anode as a negative electrode as defined in the domain of chemistry (Figure 8), such as radiology. Figures 8 and 9 show **anode** (Chemistry) (DCM – galvanic cells) and its relation to the definition of **X–ray tube** (Physics).


```
<termUnit id="1269" layer="11_01">
    <langSet xml:lang='hr-HR'>
        <term type="preferredTerm">rendgenska cijev</term>
        <definition>naprava za dobivanje rendgenskoga zračenja
        u kojoj se elektroni isijani iz užarene katode ubrzavaju
        visokim naponom u vakuumu prema metalnoj
        <target interLink="21_03" id="3670" term="anoda">anodi</target>
        u kojoj pobuđuju atome na isijavanje kontinuiranoga i
        karakterističnoga rendgenskog zračenja</definicija>
    </langSet>
    <langSet xml:lang='en-EN'>
        <term>X-ray tube</term>
    </langSet>
</termUnit>
```


Figure 9. Example of a cross–domain relation[21]

20 Translation: “definition” negatively charged particle.

21 Translation: “definition” device for obtaining X–rays in which the electrons radiated from the glowing cathode are accelerated by high–voltage in vacuum towards the metal anode in which they excite the atoms for the continuous radiation and the characteristic X–ray radiation.

Furthermore we implemented a 'type' value that can be used for multiple entries: altCat. This value indicates an alternative though compatible categorization system, whereas diffCat indicates a different incompatible categorization system. In the case of **anode**, there are two relations that are typed as alternative. They designate **anodes** defined by reaction type and by electrical potential (Table 2). A third one is related to **anode**, when it is defined as a 'positive charged electrode'.

A simulation of search results

Even though our research is still in a preliminary stage, we have simulated search results. Although the main objective was to design a terminological model that would provide a more realistic description of conceptual structures in specialized languages, it is also necessary is to create a product that will be useful for users. Since the users of Struna have a wide range of profiles, we have conceived an interface that presents query results in a simple but meaningful form.

By combining the grouping of concepts in DCM–based layers and formalizing intra and inter–domain relations as well as relations to the general language module, terminological units were configured in a well–structured conceptual network that allows users to choose the most relevant semantic environment.

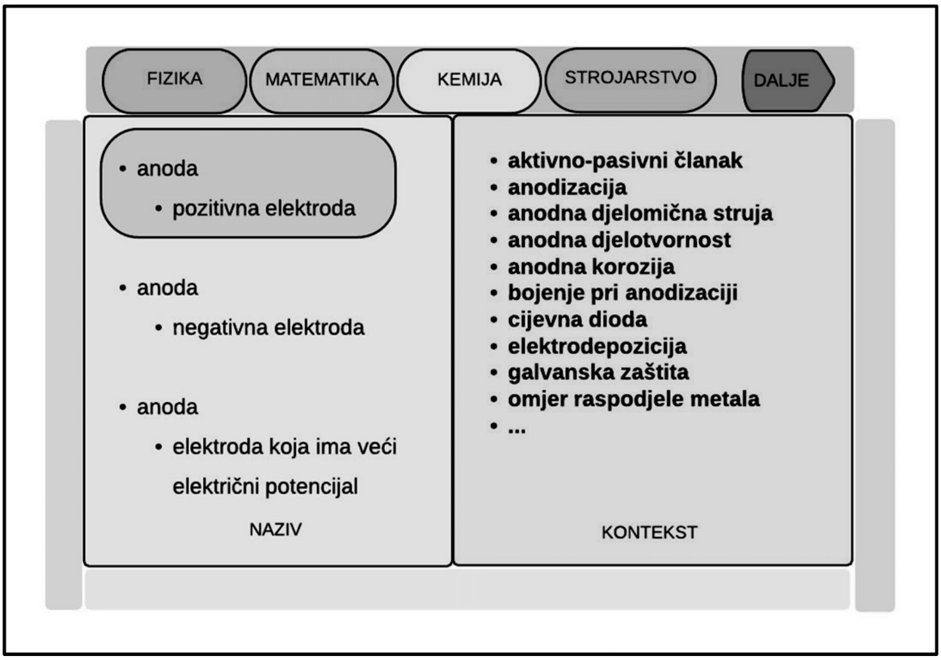

Figure 10. Simulation of results – anode/positive electrode[22]

Figures 10–12 show the simulation of one of the ways that the results for **anode** could be presented. The screen is divided in two columns with a domain bar at the top. The list of results in the left column contains a term and a definition. The right column is composed of a list of terms that corresponds to the semantic context (DCM) evoked by the specialized knowledge unit. The domain bar at the top of the screen lists the domains relevant for the DCM.

22 Translation: 1. "domain bar" physics; mathematics, chemistry, engineering; next, 2. "term column" anode – positive electrode; negative electrode; electrode with higher electrical potential, 3. "context column": active–passive cell; anodizing; anodic partial current; anode efficiency; anode corrosion; color anodizing; diode valve; electrodeposition; galvanic protection; metal distribution ratio

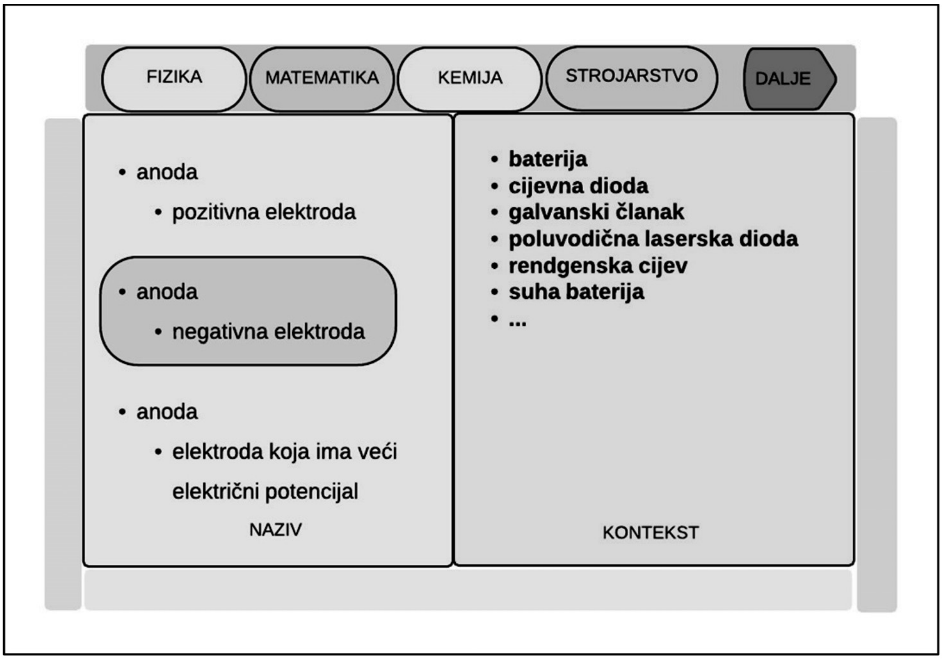

Figure 11. Simulation of results – anode/negative electrode[23]

By selecting a different term in the results column, the user activates a different DCM and the content of the context column is repopulated accordingly. From there, the user can choose whether to select a full view from a term list column or browse further by selecting one of the terms in the context column.

Although this model needs to be further developed, the results so far are promising. Our intention is to design a complementary extension that could be superimposed on the existing structure of the Struna term base. This would first allow us to test the solutions without interrupting the current workflow. Secondly, instead of developing a new database and reorganizing all of the domains, we would be able to incorporate the terminological units that are already included in it. Finally, the end user would be able to choose between a classical presentation and a DCM–based presentation of the data.

4. Conclusion

The evidence for both functional and structural differences between categorization apparatuses in general language and Language for Special Purposes (specialized domains) can be found in various research studies. Some of them are directly aimed at investigation of nature of cognition of "speakers of

23 Translation: "context" cell; diode valve, galvanic cell; semiconductor laser diode; X–ray tube; dry cell.

specialized languages" (domain experts) (Shafto and Coley 2003; Murphy and Wright 1984; Faber et al. 2014; Nahod 2015) whereas others provide indirect evidence that stems from general human cognition and general language. This study was originally motivated by problems in describing conceptual structures and relations in Croatian national multi–domain term base, Struna. Our research includes three closely related components: theoretical, empirical and practical. The results presented in this paper primarily correspond to the theoretical and practical implementations of our findings.

In this paper, we have proposed a new paradigm for processing and describing conceptual structures in specialized languages – Domain–specific Cognitive Models (DCMs). DCMs emerged as a result of the comprehensive analysis of terminological units in Struna, recent research directly related to this topic, and our own empirical study of the difference in categorization strategies between experts of the different domains and novices (Nahod 2015; Nahod and Vukša 2014). We have shown that Domain–specific Cognitive Models provide a firm foundation for both theoretical and practical processing and implementation. Practically speaking, our research goal was to find an optimal convergence point of theoretical and empirical findings, on the one hand, and the needs of end users, on the other.

It is our assertion that this multifaceted approach will lead to a highly functional model that will ensure enough flexibility to encompass specific variations both in a particular specialized language as well as between different ones. Technically speaking, our model has the advantage of being relatively easy to implement as an overlayer on the existing term base structure. This flexible model for describing categorization variations will enhance and improve the end users' experience and, more importantly, will achieve a more accurate representation of conceptual structures of processed SLs. It is to be expected that our proposal will result in term databases that accommodate more contextually–dependent knowledge and thus help the end user to better understand multidimensional concepts such as space. More specifically, users will be able to better understand why such concepts cannot be and are not perceived and defined in the same way in two or more different specialized languages.

The findings reported in this paper constitute the starting point for the second stage of our research of the nature of categorization in Language for Special Purposes, and the development of a DCM–based terminological database.

References

Bergovec Marina, Siniša Runjaić (2012). Harmonization of Multiple Entries in the Terminology Database Struna (Croatian Special Field Terminology). Guadalupe Aguado de Cea et al. eds. *Proceedings of the 10th Terminology and Knowledge Engineering Conference* (TKE 2012). 231–241.

Bergovec Marina, Siniša Runjaić (2015). Teorijske dvojbe i mogućnosti usklađivanja višestrukih terminoloških zapisa u Struni. Bratanić Maja, Ivana Brač, Boris Pritchard, eds. *Od Šuleka do Schengena*. Zagreb: Institut za hrvatski jezik i jezikoslovlje, 237–248.

Bratanić Maja, Maja Lončar (2015) Mit o uskladivanju nazivlja Europske unije iz nacionalne i europske perspektive. Bratanić Maja, Ivana Brač, Boris Pritchard, eds. *Od Šuleka do Schengena*. Zagreb: Institut za hrvatski jezik i jezikoslovlje, 223–236.

Bratanić, Maja, Ana Ostroški (2013). The Croatian National Termbank STRUNA: A New Platform for Terminological Work. *Collegium Antropologicum* 37.3: 677–683.

Capitani Erminio, Marcella Laiacona, Bradford Mahon, Alfonso Caramazza (2003). What Are the Facts of Semantic Category–Specific Deficits? A Critical Review of the Clinical Evidence. *Cognitive Neuropsychology* 20.3: 213–261.

Caramazza Alfonso, J R Shelton (1998) Domain–Specific Knowledge Systems in the Brain the Animate–Inanimate Distinction. *Journal of Cognitive Neuroscience* 10.1: 1–34.

Faber Pamela (2009). The Cognitive Shift In Terminology and Specialized Translation. *MonTI* 1: 107–134.

Faber Pamela, Carlos Márquez Linares, Miguel Vega Expósito (2005). Framing Terminology: A Process–Oriented Approach. *Meta: Journal Des Traducteurs* 50.4.

Faber Pamela, Silvia Montero Martínez, María Rosa Castro Prieto, José Senso Ruiz, Juan Antonio Prieto Velasco, Pilar León Araúz, Carlos Márquez Linares, Miguel Vega Expósito (2006). Process–Oriented Terminology Management in the Domain of Coastal Engineering. *Terminology* 12.2: 189–213.

Faber Pamela, Juan Verdejo, Pilar Leon, Arianne Reimerink, Gloria Guzman (2014). Neural Substrates of Specialized Knowledge Representation: An fMRI Study. *Revue française de linguistique appliquée,* Vol. XIX: 1–23.

Felber Helmut (1984). *Terminology Manual*. Vienna: Infoterm.

Fillmore Charles J. (1985). Frames and the Semantics of Understanding. *Quaderni Di Semantica* 6.2: 222–254.

Hanson Norwood R. (1961). *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science.* Cambridge: Cambridge University Press.

Hesse Mary (1963). *Models and Analogies in Science.* London: Sheed & Ward.

Hoffman Aaron B., Harlan D. Harris, Gregory L. Murphy (2008). Prior Knowledge Enhances the Category Dimensionality Effect. *Memory & Cognition* 36.2: 256–270.

Kaplan AS, Gregory L. Murphy (2000). Category Learning with Minimal Prior Knowledge. *Journal of Experimental Psychology* 26.4: 829–846.

Kuhn Thomas (1970). *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press.

Lakoff George (1987). *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press.

León Araúz Pilar, Pamela Faber, Silvia Montero Martínez (2012). Specialized Language Semantics. Faber Pamea. ed. *A Cognitive Linguistics View of Terminology and Specialized Language.* Berlin/Boston: De Gruyter Mouton, 95–175

Lin Emilie L., Gregory L. Murphy, EJ Shoben (1997). The Effects of Prior Processing Episodes on Basic Level Superiority. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology:* 37–41.

Mahon Bradford Z, Alfonso Caramazza (2008). A Critical Look at the Embodied Cognition Hypothesis and a New Proposal for Grounding Conceptual Content. *Journal of Physiology* 102.1–3: 59–70.

Mahut H., M. Moss (1984). Consolidation of Memory: The Hippocampus Revisited. Squire L. R., N. Butters. eds. *Neuropsychology of Memory*. New York: Guilford Press, 297–325.

Martín Antonio San, Pilar León–Araúz (2012). Flexible Terminological Definitions and Conceptual Frames. Seppälä Selja, Alan Ruttenberg, eds. *CBO 2013 Workshops Proceedings*. 1–7.

Mishkin M. (1978). Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. *Nature 273:* 297–298.

Mishkin M. (1982). A Memory System in the Monkey. *Philosophical Transactions of the Royal Society B. 298 (Biol.):* 85–92.

Moss H. E., L. K. Tyler, F. Jennings (1997). When Leopards Lose Their Spots: Knowledge of Visual Properties in Category–Specific Deficits for Living Things. *Cognitive Neuropsychology* 14.6: 901–950.

Murphy Gregory L. (2002). *The Big Book of Concepts*. Cambridge, MA: MIT Press.

Murphy Gregory L., Jack C. Wright (1984). Changes in Conceptual Structure with Expertise: Differences between Real–World Experts and Novices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10.1: 144–155.

Nahod Bruno (2015). Brak čestice i prostora: Sociokognitivna poredbena analiza pojmovnih struktura strukovnih jezika fizike i antropologije. Bratanić Maja, Ivana Brač, Boris Pritchard, eds. *Od Šuleka do Schengena*. Zagreb: Institut za hrvatski jezik i jezikoslovlje, 169–196.

Nahod Bruno, Perina Vukša (2014). On Problems in Defining Abstract and Metaphysical Concepts – Emergence of a New Model. *Collegium Antropologicum* 38 (Suppl. 2): 181–190.

Palmeri TJ, C Blalock (2000). The Role of Background Knowledge in Speeded Perceptual Categorization. *Cognition* 77.2: B45–57.

Palmeri TJ, Celina Blalock (1999). The Time–Course of the Use of Background Knowledge in Perceptual Categorization. *Cognitive Science.* http://catlab.psy.vanderbilt.edu/wp–content/uploads/File/Palmeri–CogSci1999.pdf.

Reber Paul J., Darren R. Gitelman, Todd B. Parrish, M. Marsel Mesulam (2003). Dissociating Explicit and Implicit Category Knowledge with fMRI. *Journal of Cognitive Neuroscience* 15.4: 574–583.

Rosch Eleanor, Barbara B. Lloyd, eds. (1978). *Cognition and Categorization. Principles of Categorization*. New York: John Wiley & Sons.

Rosch, Eleanor, Carolyn B. Mervis (1975). Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 605: 573–605.

Shafto Patrick, John D. Coley (2003). Development of Categorization and Reasoning in the Natural World: Novices to Experts, Naive Similarity to Ecological Knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29.4: 641–649.

Squire L. R., S. Zola–Morgan (1983). The Neurology of Memory: The Case for Correspondence between the Findings for Human and Non–Human Primate. Deutsch J. Anthony, ed. *The Physiological Basis of Memory.* New York: Academic Press, 199–268.

Stojanov Tomislav, Vedran Cindrić, Bruno Nahod, Milica Mihaljević (2009). E–Struna: baza podataka i web–sučelje hrvatskoga strukovnog nazivlja. http://struna.ihjj.hr/.

Struna. 2012. “Struna.” http://struna.ihjj.hr/ (Accessed June 2015).

Taylor H.A., I. Drugi (2001). Is the Donut in Front of the Car? An Electrophysiological Study Examining Spatial Reference Frame Processing. *Canadian Journal of Experimental Psychology* 55.2: 175–84.

Temmerman Rita (2000). *Towards New Ways of Terminology Description. The Sociocognitive Approach.* Amsterdam–Philadelphia: John Benjamins Publishing Company.

Wittgenstein Ludwig (1953). *Philosophical Investigations*. New York: Macmillan.

Wüster, Eugen (1979). *Einführung in Die Allgemeine Terminologielehre und terminologische Lexikographie.* Wien/New York: Springer.

Standards and guidelines

UNESCO Guidelines for terminology policies: formulating and implementing terminology policy in language communities. Paris: UNESCO, 2005.

ISO 704:2009(E) Terminology work – Principles and methods. Geneva: International Organization for standardization, 2009.

ISO 860:2007(E) Terminology work – Harmonization of concepts and terms. Geneva: International Organization for standardization, 2007.

ISO 1087–1:2000(E/F) Terminology Work – Vocabulary – Part 1: Theory and Application. Geneva: International Organization for standardization, 2000.

ISO 1087–2:2000(E/F) Terminology Work – Vocabulary – Part 2: Computer Applications. Geneva: International Organization for standardization, 2000.

HRN ISO 5127:2008(HR/E) Informacije i dokumentacija – terminološki rječnik. Zagreb: Hrvatski zavod za norme, 2008.

ISO 10241–1:2009(E) Terminological entries in standards – Part 1: General requirements and examples of presentation. Geneva: International Organization for standardization, 2009.

ISO 10241–2:2008(E) Terminological entries in standards – Part 2: Introduction of international terminology standards into different environments. Geneva: International Organization for standardization, 2008.

ISO 12200:1999(E) Computer applications in terminology – Machine–readable terminology interchange format (MARTIF) – Negotiated interchange. Geneva: International Organization for standardization 1999.

ISO 12615:2004(E) Bibliographic references and source identifiers for terminology work. Geneva: International Organization for standardization, 2004.

ISO 12616:2002(E) Translation–oriented terminography. Geneva: International Organization for standardization, 2002.

ISO 12620:2009(E) Terminology and other language and content resources — Specification of data categories and management of a Data Category Registry for language resources. Geneva: International Organization for standardization, 2009.

ISO 15188:2001(E) Project management guidelines for terminology standardization. Geneva: International Organization for standardization, 2001.

ISO 16642:2003(E) Computer applications in terminology – Terminological markup framework. Geneva: International Organization for standardization, 2003.

ISO 22128:2008(E) Terminology products and services – Overview and guidance. Geneva: International Organization for standardization, 2008.

ISO TR 22134:2007(E) Practical guidelines for socioterminology. Geneva: International Organization for standardization, 2008.

ISO 24610–1:2006(E) Language resource management – Feature structures – Part 1: Feature structure representation. Geneva: International Organization for standardization, 2006.

ISO 24610–2:2011(E) Language resource management – Feature structures – Part 2: Feature system declaration. Geneva: International Organization for standardization, 2011.

ISO 24611:2008(E) Language resource management – Morpho–syntactic annotation framework. Geneva: International Organization for standardization, 2008.

ISO 24613:2008(E) Language resource management – Lexical markup framework (LMF). Geneva: International Organization for standardization, 2008.

ISO 24614–1:2010(E) Language resource management – Word segmentation of written texts – Part 1: Basic concepts and general principles. Geneva: International Organization for standardization, 2010.

ISO 24614–2:2011(E) Language resource management – Word segmentation of written texts – Part 2: Word segmentation for Chinese, Japanese and Korean. Geneva: International Organization for standardization, 2011.

ISO 24615:2010(E) Language resource management — Syntactic annotation framework (SynAF). Geneva: International Organization for standardization, 2010.

ISO 24615:2011(E) Language resource management — Persistent identification and sustainable access (PISA). Geneva: International Organization for standardization, 2011.

ISO 30042:2008(E) Systems to manage terminology, knowledge and content – TermBase eXchange (TBX). Geneva: International Organization for standardization, 2008.

Domenski specifični kognitivni modeli u višedomenskoj terminološkoj bazi

Suvremeni pristupi terminološkoj teoriji temelje se na premisama kognitivne i sociokognitivne lingvistike. Iako se ti pristupi kreću od općih terminoloških okvira do više praktičnih modela koji se temelje na određenoj kognitivnoj teoriji, ni jedan od njih ne predviđa kako bi se primijenio u višedomenskoj terminološkoj bazi poput Strune. Struna (Stojanov i dr. 2009, Struna 2012) je nacionalna baza podataka hrvatskoga strukovnog nazivlja. Budući da su u Struni obrađena mnoga nazivlja iz različitih područja, nameće se potreba za razvojem novoga sociokognitivnog modela koji će uključivati premise i suvremenih terminoloških pristupa i pristupa kognitivne sociolingvistike. Takav bi model trebao omogućiti praktičan okvir za višedomensku terminološku bazu podataka.

Key words: sociocognitive linguistics, terminology, Struna
Ključne riječi: sociokognitivna lingvistika, terminologija, Struna