

# CORPUS-BASED APPROACH IN COGNITIVE SEMANTICS STUDIES

## A LINGUÍSTICA DE CORPUS APLICADA À SEMÂNTICA COGNITIVA\*

Anastasiia Belyaeva

ZAPORIZHZHYA NATIONAL UNIVERSITY, UKRAINE  
CEHUM / UNIVERSIDADE DO MINHO, PORTUGAL  
anabelwhite@gmail.com

Cognitive Semantics studies can provide the connection between linguistic and conceptual models of the world. Language is seen as the means of accessing mental processes in speaker's mind and, more broadly, the worldview and culture of a language community. However, research methods used in Cognitive Semantics studies rely on introspection, and linguists call for incorporation of more objective empirical quantitative methods. These requirements can be met by using corpora. Corpus-based studies involve looking at frequency and usage patterns to make generalizations about norms of language use. Taking up this approach, cognitive linguists would be able to look into non-elicited language data on large scale to reach conclusions about cognitive patterns of a language community. This paper describes how corpus-based approach can be used in Cognitive Semantics research.

**Keywords:** concept, Cognitive Semantics, corpus, frequency, collocation

A semântica cognitiva pode fornecer o elo de ligação entre os modelos linguísticos e conceituais do mundo. A língua é uma via de acesso aos processos mentais que se desenrolam na mente do falante e, mais genericamente, à visão do mundo e à cultura de uma comunidade linguística. No entanto, os métodos de investigação da semântica cognitiva baseiam-se na introspeção, embora se reconheça a necessidade de aplicar métodos quantitativos empíricos mais objetivos. A solução poderá residir na utilização de *corpora*. Os estudos baseados em *corpora* analisam padrões de frequência e de uso para fazer generalizações sobre as normas de uso da língua. A adotar esta abordagem, os linguistas cognitivos passariam a dispor de uma grande quantidade de dados não-elicítados, que lhes permitiria identificar os padrões cognitivos de uma comunidade linguística. Este artigo explica de que forma a metodologia da linguística de *corpus* pode ser aplicada à investigação em semântica cognitiva.

**Palavras-chave:** conceito, semântica cognitiva, *corpus*, frequência, colocação

---

\* Agradecemos aos revisores anónimos os seus comentários relevantes, que permitiram corrigir e precisar aspetos deste texto. Quaisquer erros e imprecisões remanescentes são, naturalmente, da nossa responsabilidade.

## 0. Introduction

Cognitive Semantics assumes that the meaning of a language unit is rooted in the conceptual system that encodes encyclopaedic knowledge of the world (Langacker, 2001). The reason for the increased interest in Cognitive Semantics is its ability to provide the connection between linguistic and conceptual models of the world. Thus, language is increasingly seen as the means of accessing mental processes in speaker's mind and, more broadly, the worldview and culture of a language community.

Even though Cognitive Semantics offers no unified approach to the explanation of language phenomena, most of the methods used involve qualitative analysis and introspection. Even though such methods bring valuable results, they can hardly be suitable for making generalizations about conceptual structures characteristic of the whole language community. Currently more linguists call for incorporation of empirical quantitative methods, claiming that such analysis would make the results of Cognitive Linguistics research more objective (Gries, 2014; Geeraerts, 2006: 21–300). These requirements can be met by using corpora, i.e. carrying out quantitative analysis in a collection of written texts or transcribed utterances. This approach involves looking at frequency and usage patterns to make generalizations about norms of language use. Taking up corpus-based approach, cognitive linguists would be able to look into non-elicited language data on large scale to reach conclusions about cognitive patterns of a language community.

However, it can be argued that Corpus Linguistics research mostly relies on the analysis of particular words or word forms, while concepts and conceptual structures can be expressed by several linguistic items (Stefanowitsch, 2006: 64). This paper describes how corpus-based approach can be used in Cognitive Semantics research. The objective of the article is reviewing procedures of Cognitive Semantics analysis in order to establish Corpus Linguistics techniques that can be used in conceptual research. The study begins with a brief overview of the ways, in which human experience and knowledge can be encoded in a conceptual system and subsequently expressed in a language. The paper proceeds to determine how meaning is treated in Cognitive Semantics studies. Finally, the study attempts to outline the corpus-based procedures linguists can use to study Cognitive Semantics.

## 1. Concepts as Units of Cognition

In Cognitive studies human experience, feelings, sensations, and knowledge gained about the objective reality are the basis for the formation of mental categories. Categorization can be listed among general cognitive abilities along with abstraction and schematization (Langacker, 1999: 25). It is the process of assigning an object or phenomenon to a category by similarity; it is based on hyponymy - hyperonymy relations and deals with objects or phenomena in the real world (Margolis, 1994; Teixeira, 2001: 66). Categories, in their turn, become subject to mental transformation or cognitive-semantic conversion. This mental process systematizes categories and organizes them into concepts; the latter can be regarded as mental structures derived from reality and based on sensations (Teixeira, 2001: 8). Delineation between real world and mental phenomena is further stressed in Eastern Europe where Cognitive Studies use different terms to refer to cognitive representation of abstract notions (term 'concept') and the notions in question. In Cognitive language studies researchers use umbrella term 'concept' (Степанов, 2001: 40-76) or more specific research units, such as 'linguocultureme (mental unit combining linguistic and culture-specific information)' (Воркачев, 2007; Демьянков, 2001), 'mythologeme (unit of thinking based on mythological beliefs encoded by language)' (Базылев, 2000)), 'logoe pistheme (epistemic mental unit that can be accessed through language)' (Бурвикова, 2001; Верещагин, 1999: 7). Nevertheless, studies that adopt the abovementioned terms deal with the analysis of mental representation resulting from categorization of objects and phenomena of the world. The difference in the coined terms stems from scholars' focus on cultural relevance of the concept ('mythologeme' and 'linguocultureme') or its rational component ('logoe pistheme'). The present study adopts the broad term 'concept'.

Views on relationship between human knowledge and its presentation in a language vary from approaches within the framework of traditional semantics to purely mental understanding of a concept as a unit that is not expressed by language (Кубрякова, 1991: 85; Petersen, 2007: 426). Nowadays linguists use one of the two major approaches to concept analysis: cognitive or linguo-cultural. The first approach treats concepts as units encoding memories and experiences of a person. In this view concepts are treated as primary nonverbal units of thinking. However, these mental units can be expressed by language means. However, language verbalizes only part of conceptual structure, the elements that are the most signifi-

cant and relevant for the given language community. Therefore linguists analyse what components of concept(s) are reflected in the language in order to access the conceptual structure of the whole language community (Стернин, 1998: 24-27).

Linguo-cultural approach treats concepts as multidimensional mental units that are verbalized in language and reflect specific features of culture, i.e. they reflect culturally relevant experiences. In this view, concepts are not units of person's memory, but elements of culture and collective memory of a language community (Воркачев, 2007; Карасик, 2001: 4; Степанов, 2001: 41). Therefore, concepts do not only represent real world phenomena or abstract notions, but also contain information about speakers' (collective) experience with the phenomenon in question, i.e. what the phenomenon is associated with, how it is evaluated, whether it is important for the language community. However, no matter which of the two approaches linguists favour, they rely on the language means in order to study concepts.

In cognitive studies concepts are seen as multidimensional structures. Being mental units, they include not only the knowledge of the real world phenomena, but also broader cultural relevance of the notions or objects they represent. Concepts are believed to be linked to primary bodily sensations and perceptions that become mental structures as a result of abstraction. Being units of thinking, concepts are organized according to the way different elements of reality (objects, events) influence human perception. Thus, through repeated exposure associative relationship between a phenomenon of the world and the impact it has on perceptual or somatic sphere of an individual is formed. This relationship results in the creation of a generalized image, mental "footprint" of the qualities and properties of the element of the physical world that becomes a concept (Чекулай, 2006: 16). Concepts emerge as rational (notional) and perceptive (sensations-based) components of human experience are combined. Rational (logical) component of the concept is the result of the process of conceptualization; it reflects the structure and characteristics of the relevant phenomena and notions (Воркачев, 2001: 56). From the realm of individual's mental sphere concepts are transferred to the collective realm of culture. There mental units undergo evaluation and acquire a set of associations. Evaluation of the concept constitutes interpretive field or periphery of a concept. It includes such elements as figurative component (conceptual metaphor), and axiological (value) component. Therefore, concepts can be seen as complex units made of different layers. Some of them are crucial to the formation of a concept

and are therefore its core; others are results of interpretation or assessment of the notion. Core of the concept contains the most significant information about the results of the speakers' conceptualization. It combines components that are formed as a result of human cognitive and sensory perception of the world, in other words rational (logical) and perceptive layers. The combination of the rational and perceptual conceptual components is possible through the mental operation of image schema creation (Johnson, 1987: 29; Turner, 1991). The latter "help to explain how our intrinsically embodied mind can at the same time be capable of abstract thought. As patterns of sensory-motor experience, image schemas play a crucial role in emergence of meaning and in our ability to engage in abstract conceptualization and reasoning that is grounded in our bodily engagement with our environment" (Johnson, 2005: 15). Image schema is an embodiment of prelinguistic experience. It motivates conceptual mappings and operates as a dynamic pattern of relating perceptual interactions or motor programs structuring human experience. Without such integrating procedure human experience would be chaotic and incomprehensible (Johnson, 1987: xix; Kövecses, 2006: 207). Image schemas make it possible to rely on the structure of sensory and motor operations to understand abstract concepts and draw inferences about them. They are repeated patterns of sensory-motor experience of the individual (Johnson, 1987: XIV).

Analysis of concepts is a series of procedures to clarify characteristics of concepts as units of cognition and describe their relations within system. Currently concept analysis focuses on concept system in terminological analysis and nursing science (Nuopponen, 2010; Walker, 1994). In concept analysis the structure of concepts is presented as basic, structural, origination, development, activity, transmission, causation or dependency models. Basic model helps to establish concept's place in generic (logical) concept system, structural model reveals spatial relations between referents of the concepts. Origination and developmental models are used to establish relations between the core concept and phenomena that either are connected to its origin or reflect stages in its development. Activity and transmission models connect the core concept (representing activity) and phenomena involved in this activity. Causation and dependency models establish relations between cause and effect, patient affected by the core concept, types of effect. Components from all models can be used at different stages of research depending on research objectives. They can be integrated in a single concept map model (Nuopponen, 2010: 10-12). Although existing models were not elaborated within the context of cognitive semantics, they

stress that the analysis of a concept is basically the analysis of its lexical representation and its usage, an examination of how the word(s) that denote concept is related to other units of language (Walker, 1994: 38-41).

## **2. Cognitive Semantics and Study of Concepts**

Cognitive Semantics assumes that language reflects conceptual system and can provide access to conceptual structuring of the world. One of the guiding principles in Cognitive Semantics equates semantic and conceptual structures because language deals with mental representations of the objective reality in speakers' mind and not directly with the real world phenomena (Evans 2007, p. 6). In the central postulate of Cognitive Semantics meanings are seen as motivated mappings from conceptualization to expression (Sinha, 1999: 231). However, scholars do not perceive the two structures (cognitive and semantic) as identical. Even though language conceptualization is rooted in pre-linguistic or non-linguistic schematization, it forms only a subset of the concepts as mental units. In this view conceptualization is not necessarily supported by language means; even basic cognitive operations of prototype formation are seen as being "compatible with the language", but not necessarily language-supported (Allwood, 1999: 4-6). Scholars see meanings as dynamic entities formed in the discourse practise and not as equivalents of concepts and schemata (Sinha, 1999: 233-237).

Conceptual structure is believed to give access to a vaster repository of knowledge derived from sensory-motor and proprioception systems as well as subjective experience. This knowledge is encyclopaedic in its nature and cannot be fully reflected in language semantics. This approach has led some scholars, namely V. Evans, to distinguish lexical concepts from cognitive models (Evans, 2007). In this view, lexical concepts that encode linguistic knowledge are reflected in semantic structure, while conceptual structure is non-linguistic in its nature and is reflected in conceptual system. The latter derives from human experience in the world and its perceptual nature does not allow it to be encoded in language. However, conceptual structure can take informational form that can be directly represented in a language. Therefore lexical concepts can be seen as related to non-linguistic conceptual structures and are treated as the form the latter take to be directly encoded in language (Evans, 2009: 30). This approach to Cognitive Semantics analysis assumes that lexical form can give access to the inventory of lexical concepts it expresses (Evans, 2010). Therefore,

two main trends in Cognitive Semantics research can be singled out: 1) research looking into semantics of a language unit to establish concept(s) it represents (direction: from language unit(s) to concept(s)); 2) studies aimed at establishing the language means a concept as a mental unit takes in language (direction: from concept to language unit(s)).

Analysis of concepts focuses on uncovering cognitive mechanisms that underlie the formation of knowledge about the real world phenomena and abstract notions. In the broad sense, conceptual analysis is the study of properties and characteristics of concepts and their relations to other concepts. In Cognitive Semantics conceptual analysis is the method of identifying and studying the structure of verbalized concepts, i.e. results of human cognition reflected by language means. Language semantics is rooted in the system of concepts that encodes human experience and worldview. Semantics can give a key to understanding individual language units and mechanisms of human knowledge about the world. However, study of concepts does not equal analysis of word semantics. The former tries to establish general concepts, while the latter deals with the meaning of separate words. Meaning of language units is built of concepts; therefore semantics can relate words to the world. In this view, words are mapped onto conceptual structure and polysemous words can verbalize several concepts (Kurteš, 2006: 1200).

In Cognitive Semantics there is no universal method to the study concepts; therefore scholars rely on a combination of techniques whose priority is determined by research objectives. Research procedures used by linguists can be based on: 1) methods borrowed from philosophy and anthropocentric linguistics: a) logical approach; b) ethnocentric approach; 2) knowledge processing analysis; 3) eidetic analysis. In order to study language phenomena linguists utilize logical models, such as frames and scripts (Fillmore, 1982). Ethnocentric studies use metalanguage to describe the meanings of language units and establish underlying universals represented by languages (Wierzbicka, 1985). Analysis in terms of the knowledge processing method seeks to establish how human experience is interiorized by cognition. It relies on conceptual metaphor, metonymy, and image schemas (Clausner, 1999). Eidetic analysis attempts to uncover natural categories in language; the studies are conducted within the framework of prototype theory (Rosch, 1976). Yet, in these approaches linguists rely on their own judgement and/or language intuition when they establish the connection between language units and mental structures.

Cognitive Semantics developed two approaches to the analysis of concepts: 1) cognitive interpretation; 2) concept modeling. Cognitive interpretation approach relies on the analysis of semantics of language units that name the concept under study. Firstly, it is established what language means are used to reflect the knowledge of a certain phenomenon. Those language means are regarded as name(s) of the concept; they are often retrieved through the dictionary definitions analysis. Then, scholars establish lexical compatibility of the language phenomena in question. In this approach definitions of names of the concepts under analysis are used to establish cognitive properties of the concept. Researchers try to describe the meaning of language units and then utilize the techniques of cognitive interpretation of the results to model the structure of a concept. The study of the semantic structure of the lexical units that verbalize the concept is seen as the key to the mechanisms of conceptualization (Жаботинская, 2005: 53-55; Gladkova, 2008; Łodej 2012). In this view semantic structure of name(s) of concept is mapped onto mechanisms of knowledge processing. Within the second approach concept is regarded as a multidimensional unit that reflects human perception of certain elements of reality. Different dimensions of concept as unit of cognition correspond to different layers in its structure. Like cognitive interpretation, research starts with creating inventory of language means that represent the concept. However, concept modeling focuses not only on semantics of language means, but also takes into account cultural context those means exist in. Scholars use the contexts name of the concept is used in to access information about the language community's views on the phenomenon that concept represents. This approach studies meaning of language units in broader context. Semantic analysis is regarded as a stage in conceptual analysis in the present study. Therefore, the study of concepts in Cognitive Semantics requires interpretation of associations that reflect stereotyped knowledge, beliefs, assumptions, and evaluations that are associated with the phenomenon concept represents.

Within the context of cognitive semantics concept modeling is singling out a number of components (modes) in the concept under analysis. The number of components or layers in verbalized concepts varies depending from research objectives. Characteristics of concepts as units of cognition reviewed in the previous section of the paper can be reduced to three major concept components: notional, perceptive and axiological layers. Notional component of a concept encompasses its language representation, its name. This component is established through analysis of semantic structure of



language units that verbalize concept; its characteristics as compared to other groups of concepts are also established. Perceptive component verbalizes visual, auditory, tactile, taste characteristics of objects or events reflected in human consciousness (Воркачев, 2001: 49). Cognitive semantics seeks to uncover what scenarios of embodied knowledge acquisition the name of the concept is involved in. This task is achieved by looking into collocation patterns of language units under analysis. The task of conceptual analysis is to present a concept in schematic form as a structure made of central (core) layers and the periphery. According to R. Langacker, language usage profiles the most significant parts of the content of the concept that serve as the speaker's focus of attention (Langacker, 1987: 145). Verbal explication of the actualized seme is the linguistic operation similar to profiling (Стернин, 2006: 77-78). The study of rational (logical) component and perceptive modes of concept is identification of the semes in the names of the concept that are 'profiled' in the context. The study of the structure of rational (logical) layer of the concept provides an inventory of its components that are verbal equivalents of notional components of concept. Somatic knowledge that is part of conceptualization can be studied through the analysis of conceptual metaphors which are the result of cognitive operations of correlation between the structure of the source domain and the target domain (Ruiz de Mendoza, 2005: 250). Evaluation of objects and phenomena is verbalized in the axiological mode of a concept (Бессонова, 2007: 227). This layer is analyzed by establishing the role and place of the phenomenon in culture because concepts exist as units of reasoning and perception, and the latter involves evaluation. In cognitive semantics this layer is modeled through the study of connotation as well as evaluative meaning that manifests itself in collocations that language units form.

No matter which of the approaches linguists take, they look at word usage in real contexts in order to study meaning. However, scholars should be warned not to make generalizations about underlying cognitive structures using small language samples. Researchers suggest that very often "basic cognitive concepts are postulated on the grounds of poor linguistic data and their cognitive relevance is then "proved" by finding them realized in the investigated language" (Blank, 2003: 44). By addressing larger language samples linguists would be able to formulate more feasible hypotheses about cognitive structures reflected by language(s).

### 3. Corpus Linguistics Application for Cognitive Studies

Corpus-based research in linguistics relies on using automated techniques and quantitative (statistical) methods in order to uncover patterns of language usage or determine similarity and difference between language units. Since corpus queries are largely based on the search for individual word forms, this approach does not seem immediately relevant for semantics studies. The limited applicability of corpus techniques can be grounded in the nature of semantic investigation. As G. Sampson stresses “corpora have limited relevance (though some relevance) to the study of semantics (...) because the study of semantics is not an empirical scientific discipline at all, but something more like a branch of philosophy” (Sampson, 2011: 203). However, maintaining the right balance between qualitative and quantitative models of analysis, linguists gain new insights into word meaning. Among major findings of corpus-based semantics is the notion of semantic prosody (Charteris-Black, 2004; Stubbs, 2001). The evidence of evaluative meaning that manifests itself in repeated patterns of typical word use, i.e. semantic prosody, was gained owing to qualitative analysis of word collocations retrieved from corpora. Therefore, the study of the collocations of a word is among the most widely used corpus-based research procedure in semantic analysis. Collocation can be seen as a lexical relation between words that co-occur in a natural text in a statistically significant way (Stubbs, 1995; Lewis, 2002). This method is helpful in the studies of connotative and discursive meaning of language units; it gives insight into the changing norms of language use and acceptability of collocations; it can be incorporated into error analysis in language acquisition studies. However, relevance of corpus-based techniques in Cognitive Semantics and the study of concepts is yet to be established.

Corpus-based approach can be of benefit for Cognitive Semantics studies if it develops methodology appropriate for both qualitative (creating the inventory of language means that reflect concepts) and quantitative (techniques to extract language units, their collocates, calculate their frequency) analysis. The study of the word collocations can give insight into the difference between language units that verbalize concepts, since even though “two words can be conceptually similar in their basic and even extended meanings, but the collocational patterns are usually highly language specific” (Verdauguer, 2003: 634-635). What is more, linguists assume that “(...) from habitual, i.e. typical and frequent, expressions of a language we can infer a speech community’s habitual ways of conceptualization” (Schönefeld, 2007: 298).

The starting point for establishing connection between conceptual and semantic relations can be found in J. Saeed's views on the nature of conceptual knowledge: "words are in a network of semantic links with other words and it is reasonable to assume that conceptual structures are similarly linked" (Saeed, 1997: 38). Therefore, linguists conducting corpus-based research into Cognitive Semantics turn to lexical reference systems like WordNet ([www.wordnet.princeton.edu](http://www.wordnet.princeton.edu)) (Ahrens, 2011; Davies, 2007; Fellbaum, 1998). In this resource lexical units (nouns, verbs, adjectives, adverbs) representing underlying concepts are organized into synonym sets. Having retrieved the set of lexical units that represent concept(s) under analysis, linguists can proceed to look for the collocations these words enter using corpus queries. As K. Krawczak states: "it is possible to reveal the conceptual structure behind language by analysing its contextualized recurrent use, both qualitatively and quantitatively" (2014: 446).

Nowadays, two major approaches to corpus-based studies in semantics are Latent Semantics Analysis (relying on statistical analysis of collocations) and Multifactorial Usage-Feature analysis (relying on multivariate statistics of array of language features) (Glynn, 2009). In the first approach quantitative analysis helps to determine the degree of typicality of collocation and find out what characteristics of the concept under study are important for the language community. Nowadays, Latent Semantics Analysis is used in studies of cognitive metaphors and words belonging to a specific conceptual domain. Semantic corpus-based analysis can rely on frames, key lexemes that form cognitive metaphor models, image schemas and lexical units that represent them, units of Natural Semantic Metalanguage. In cognitive-semantic studies of metaphors linguists determine what should be searched for in a corpus using lexical units associated either with: 1) a specific conceptual model, for example K. Ahrens (2011) used WordNet to identify key lexemes representing metaphoric models in US presidential speeches; 2) a specific target domain, for example emotions in research conducted by A. Stefanowitsch (2006) or physical environment and plants in the corpus-based study of metaphor in political discourse by J. Charteris-Black (2002; 2004). Image schemas as basic pre-conceptual representations of human sensory-motor experience determine the choice of verb collocations, namely posture verbs, in cognitive comparative corpus-based study of posture scenarios by D. Schönefeld (2007). Universal primary meanings encoded in Natural Semantic Metalanguage (developed by A. Wierzbicka and C. Goddard) determined the choice of lexical units in corpus-based contrastive study aimed at reconstruction of concept TOL-

ERANCE in the investigation conducted by A. Gladkova (2008). Affiliation with conceptual domain CLERGY motivated the selection of language units in diachronic corpus-based research conducted by S. Łodej (2012). In case of idioms, cognitive metaphors or concepts that are represented by several language units linguists can start by looking at a corpora sample to qualitatively analyze linguistic expressions and narrow the scope of items to be further searched for and quantitatively assessed. In this type of analysis scholars can rely on frequency of a certain language unit or word in a corpus in order to make conclusions about speakers' cognitive patterns. Once the array of language units is established, it can be analyzed using concordance and statistical corpus-research tools. Quantitative data is used to determine what collocations, and subsequently underlying mental patterns, are more typical of or acceptable in the given language community.

In Multifactorial Usage-Feature analysis a sample of data extracted from the corpora is qualitatively analyzed and manually annotated using a combination of usage and meaning criteria. Then the linguists use the sample corpus to determine the significance and behavior probability of a certain language unit, i.e. they apply quantitative techniques to the sample set of data to make predictions about language behavior (Krawczak, 2014). For example, D. Glynn used semantic frames in the research on polysemous words using corpora downloaded from the web (Glynn, 2009: 83-84). However, this technique has some limitations that make it difficult to apply to Cognitive Semantics research. Firstly, the analysis is carried out on a relatively small set of data, because Multifactorial Usage-Feature annotation is done manually, which would not completely meet the objectivity criteria. Secondly, categories for manual annotation are established prior to corpus analysis and not necessarily through analysis of language units. In current studies Multifactorial Usage-Feature annotation includes a combination of formal, semantic, and sociolinguistic features (Krawczak, 2014). Therefore, it is not suitable for research aimed at establishing an array of concepts that underlie language units. It favors only the studies that start with concept(s) and then proceed to language means used to express them.

As the concept has a complex structure, its analysis has to combine a number of methodologies. The structure of concept in cognitive semantics can be established following two major stages of analysis. During the first stage lexical units that verbalize the concept are established, cognitive properties of the concept are then defined. Concept analysis calls for the creation of "semantic inventory" of the means that language under study

uses to express certain concept. This stage relies on the study of semantic structure of words representing a concept in a language to reveal the principles according to which information is organized in mind. At this stage corpus-based model of Latent Semantics Analysis, namely quantitative study of keywords would help narrow the scope of language units that verbalize the concept. In order to give a key to understanding the mechanisms of conceptualization, this stage should involve analysis of rational (logical) component of the concept structure as well as its perceptive component. This stage also looks into figurative components of concepts represented by cognitive metaphors. Corpus-based collocation analysis would provide the foundation for the study of actualized senses in context looking into the frequent and/or typical words that the name(s) of the concept is combined with. Next procedure involves looking into axiological (value) component of concept. Corpus-based technique that would allow looking into this layer of concept structure is the analysis of collocations to study semantic prosody (see Charteris-Black, 2004; Stubbs, 2001). The outcome of the first stage is modeling a concept structure. Next step in concept system analysis is establishing the place of concept in a generic concept system and attempt to reveal its relations (coordination, subordination, and superordination) to other concepts. At this stage concept analysis needs to establish concept hierarchy. Even though a study may focus on a single concept, concepts never exist in isolation. Concept relation types are based on structures and categorizations of the world and conceptual hierarchies can be based on hierarchies of semantic relations (Nuopponen, 2010: 10). There are semasiological and onomasiological components in this stage. Cognitive semasiological analysis takes into account different parts of speech that denote a concept. Frames are used to represent different relations between lexical units that verbalize the concept (Кустова, 2004). Cognitive onomasiological analysis is an attempt to locate concepts depending on the level of their abstraction (Селиванова, 2000). Second stage in the analysis relies on Multifactorial Usage-Feature model. Once the scope of the language units that verbalize the concept under analysis is narrowed down through quantitative and qualitative collocation analysis, the most frequent and/or typical ones can be manually annotated for usage and meaning criteria in order to access broader cultural significance of the concept.

## 4. Conclusions

The conclusions drawn from the research suggest that the complex nature of human language allows it to be both a tool of world understanding and a means of storing and transferring the knowledge gained. The results of categorization of real world objects and phenomena constitute conceptual representation of the world. Since learning about objects and phenomena people use both their senses and cognitive abilities, concepts can be regarded as units combining rational and perceptual components. As parts of culture concepts also include associations and evaluation. In this view concepts can be regarded as multidimensional structures made of several layers. Concepts may be expressed by language means. Cognitive Semantics analysis of linguistic units that name concepts can give access to conceptual model of the world. This paper suggests that empirical methods characteristic of corpus-based approach can be suitable for Cognitive Semantics studies because they would provide a more feasible foundation for generalizations about worldview of a language community. Even though corpus-based approach is often used in semantics research, there is no unified view on the procedures applicable for Cognitive Semantics. Complex nature of concepts as mental units that can be represented in language only partially makes it necessary to combine quantitative corpus-based methods with well-established Cognitive Semantics procedures of language analysis. Delimitation of language units that verbalize a concept would be a starting point for further corpus-based research. In case linguists favor a different approach to concept analysis, i.e. they consider that one language unit expresses several concepts, an array of concepts that a language unit expresses is established prior to corpus queries. This stage can benefit from quantitative analysis techniques aimed at establishing keywords in corpus. Looking into collocational patterns linguists would gain insight into the place concept occupies in conceptual system of the community, understand what other notions or phenomena it is typically associated or considered alongside with. Once the quantitative analysis of collocations is interpreted, its results could be used in more sophisticated Multifactorial Usage-Feature analysis. The paper advocates the combination of the two corpus-based approaches to semantics analysis where Latent Semantics Analysis provides a more objective (quantitatively-tested) and language-based foundation for Multifactorial Usage-Feature annotation.

## References

- Базылев, В. Н. (2000). Мифологема скуки в русской культуре. In В. Н. Базылев (Ed), *RES LINGUISTICA: сб. статей к 60-летию проф. В. П. Нерознака* (pp. 130–147). М: Academia.
- Бессонова, О. Л. (2007). Особливості дієслівної вербалізації оцінних концептів в англійській мові. In *Одиниці та категорії сучасної лінгвістики*. (pp. 227-241). Донецьк : Вид. центр ДонНУ.
- Бурвикова, Н. Д., Костомаров, В. Г. (2001). Лингвокультурологические единицы в обучении русскому языку. Лингвистическое и культурное разнообразие – фактор европейского развития. *Русский язык в контексте Европейского года языков* : сб. науч. трудов, 15–21.
- Верецагин, Е. М. (1999). В поисках новых путей развития лингвострановедения: концепция речеповеденческих тактик. М.: Гос. ин-т рус. яз. им. А. С. Пушкина.
- Воркачев, С. Г. (2001). Концепт «счастье»: понятийный и образный компоненты. In *Известия РАН. Серия “Литературы и языка”*, 60, № 6, 47–58.
- Воркачев, С. Г. (2007). Лингвокультурная концептология: становление и перспективы. *Известия РАН. Серия “Литературы и языка”*, 66, 2, 13–22.
- Демьянков, В. З. (2001). Понятие и концепт в художественной литературе и в научном языке. *Вопросы филологии*, 1, 35–47.
- Карасик, В. И. (2001). Лингвокультурный концепт как единица исследования. In В. И. Карасик & Г. Г. Слышкин (Eds), *Методологические проблемы когнитивной лингвистики* (pp. 75–80). Воронеж: Изд. ВГУ.
- Кубрякова, Е. С. (1991). Об одном фрагменте концептуального анализа слова “память”. *Логический анализ языка. Культурные концепты*: сб. науч. статей, 85–91.
- Кутова, Г. И. (2004). Типы производных значений и механизмы языкового расширения. М.: Языки славянской культуры.
- Селиванова, Е. А. (2000). *Когнитивная ономазиология*. К.: Изд-во украинского филосоциолог. центра.
- Степанов, Ю. С. (2001). *Константы: словарь русской культуры : опыт исследования* (2-е изд., испр. и доп.). М.: Акад. проект.
- Стернин, И. А. (2006). Моделирование концепта в лингвоконцептологии. In *Международный конгресс по когнитивной лингвистике: сб. материалов*. Тамбов : Изд-во ТГУ им. Г. Р. Державина, 77-79.
- Стернин И. А. (1998). Национальная специфика мышления и проблема лакунарности. In И. А. Стернин (Ed), *Связи языковых единиц в системе и реализации* (pp. 22–31). Тамбов: ТГУ.

- Чекулай, И. В. (2006). Функционально-деятельностный подход к изучению принципов оценочной категоризации в современном английском языке. Белгород: Изд-во Белгородского университета.
- AHRENS, K. (2011). Examining conceptual metaphor models through lexical frequency patterns: A case study of U.S. presidential speeches. In S. Hall & H.J. Schmid (Eds), *Windows to the mind: metaphor, metonymy and conceptual blending* (pp. 167-184). Berlin/New York: Walter de Gruyter.
- ALLWOOD, J. (1999). Semantics as Meaning Determination with Semantic-Epistemic Operations. In J. Allwood & P. Gärdenfors (Eds), *Cognitive Semantics: meaning and cognition* (pp. 1-17). John Benjamins.
- BLANK, A. (2003). Words and Concepts in Time: towards Diachronic Cognitive Onomasiology. In R. Eckardt & K. von Heusinger & Ch. De Schwarze (Eds), *Words in Time: Diachronic Semantics from Different Points of View* (pp. 37–65). Walter de Gruyter.
- CHARTERIS-BLACK, J. (2004). *Corpus Approaches to Critical Metaphor Analysis*. Palgrave Macmillan.
- CHARTERIS-BLACK, J. (2002). Why “An Angel Rides in the Whirlwind and Directs the Storm”: A Corpusbased Comparative Study of Metaphor in British and American Political Discourse. In K. Aijmer & B. Altenberg (Eds), *Advances in Corpus Linguistics* (pp. 133–150). Amsterdam: Rodopi.
- CLAUSNER T. C., Croft, W. (1999). Domains and image schemas. *Cognitive Linguistics*, 10, Vol. 1, 1–31.
- DAVIES, M. (2007). Semantically-based Queries with a Joint BNC/WordNet Databases. In R. Facchinetti (Ed), *Corpus Linguistics: 25 years on* (pp. 149-167). Amsterdam, New York: Rodopi.
- EVANS, V. (2007). The Cognitive Linguistics Enterprise: An Overview In V. Evans & B. K. Bergen & J. Zinken (Eds), *The Cognitive Linguistics Reader* (pp. 1–36). London: Equinox.
- EVANS, V. (2009). Semantic representation in LCCM Theory. In V. Evans & S. Pourcel (Eds), *New Directions in Cognitive Linguistics* (pp. 27–55). John Benjamins.
- FELLBAUM, Ch. (Ed) (1998). *WordNet. An electronic lexical database*. Cambridge, MA: MIT Press.
- FILLMORE, Ch. J. (1982). Frame semantics. *Linguistics in the morning calm: Selected papers from the SICOL* (pp. 11-38). Seoul: Hanshin.
- GEERAERTS, D. (2006). Methodology in Cognitive Linguistics In G. Kristiansen & M. Achard & R. Dirven & F. J. Ruiz de Mendoza Ibáñez (Eds), *Cognitive Linguistics: Current Applications and Future Perspectives* (pp. 21–49). Berlin: Mouton de Gruyter.
- GLADKOVA, A. (2008). Tolerance: New and traditional values in Russian in comparison with English in Cross-linguistic semantics. In C. Goddard (Ed), *Explications, metalanguage, grammar* (pp. 301–329). John Benjamins.



- GLYNN, D. (2009). Polysemy, syntax, and variation. A usage-based method for Cognitive Semantics. In V. Evans & S. Pourcel (Eds), *New Directions in Cognitive Linguistics* (pp. 77–104). John Benjamins.
- GRIES, S. Th. (2014). Corpus and quantitative methods. In J. Taylor & J. Littlemore (Eds), *The Bloomsbury Companion to Cognitive Linguistics* (pp. 279–300). London & New York: Bloomsbury.
- JOHNSON, M. (1987). *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: Univ. of Chicago Press.
- JOHNSON, M. (2005). The philosophical significance of image schemas. In B. Hampe (Ed) *From perception to meaning: Image schemas in cognitive linguistics* (pp. 15–33). Berlin/New York: Mouton de Gruyter.
- KÖVECSES, Z. (2006). *Language, mind and culture: a practical*. Oxford: Oxford University Press.
- KRAWCZAK, K. (2014). Shame, embarrassment and guilt: corpus evidence for the cross-cultural structure of social emotions. *Poznań Studies in Contemporary Linguistics*, 50(4), 441–475.
- LANGACKER, R. W. (1999). Assessing the cognitive linguistics enterprise. In Th. Johannes & G. Redeker (Eds), *Cognitive Linguistics, Foundations, Scope, and Methodology* (pp. 13–59). Berlin, New York: Mouton de Gruyter.
- LANGACKER, R. W. (2001). Discourse in Cognitive Grammar. *Cognitive Linguistics*, 12 (12-2), 143–188.
- LANGACKER, R. W. (1987). *Foundations of cognitive grammar*: [in 2 vol.]. Vol. 1: *Theoretical Prerequisites*. Stanford: Stanford University Press.
- ŁODEJ, S. (2012). Concept-driven semasiology and onomasiology of CLERGY: focus on the lexicogenesis of POPE, BISHOP and PRIEST. In M. Markus & Y. Iyeiri & R. Heuberger & E. Chamson (Eds), *Middle and Modern English Corpus Linguistics* (pp. 93–108). Amsterdam: John Benjamins.
- MARGOLIS, E. (1994). A reassessment of the shift from the classical theory of concepts to prototype theory. *Cognition*, 51, 73–89.
- NUOPPONEN, A. (2010). Methods of concept analysis – towards systematic concept analysis. In *The LSP Journal: Language for special purposes, professional communication, knowledge management and cognition, Volume 1(1)*, 5–14.
- PETERSEN, W. (2007). Conceptual fingerprints: lexical decomposition by means of frames – a neuro-cognitive model. In W. Petersen & M. Werning Priss & S. Polovina & R. Hill (Eds), *Conceptual structures: Knowledge architectures for smart applications* (pp. 415–428). Heidelberg: Springer-Verlag.
- ROSCH, E., Mervis, B. ., Gray, W. D. ., Johnson, D. M. ., Bayes-Braem, P., (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382–439.

- RUIZ DE MENDOZA, F. J. Conceptual Interaction, cognitive operations, and projection spaces. In F. J. Ruiz de Mendoza, S. M. Peña (Eds) *Cognitive Linguistics. Internal Dynamics and Interdisciplinary Interaction* (pp. 249–282). Berlin, New York: Mouton de Gruyter.
- SAMPSON, G. (2011). A two-way exchange between syntax and corpora. Interview. In V. Viana & S. Zyngier & G. Barnbrook (Eds), *Perspectives on corpus linguistics* (pp. 197-212). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- SAEED, J. (1997). *Semantics*. Oxford: Blackwell publishers Ltd.
- SINHA, Ch. (1999). Grounding, Mapping and Acts of Meaning. In Th. Janssen & G. Redeker (Eds), *Cognitive Linguistics: Foundations, Scope and Methodology* (pp. 223–255). Berlin & New York: Mouton de Gruyter.
- STEFANOWITSCH, A. (2006). Words and their metaphors: a corpus-based approach. In A. Stefanowitsch & S. Th. Gries (Eds) *Corpus-based approach to metaphor and metonymy* (pp. 63-105). Berlin: Walter de Gruyter.
- STUBBS, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell.
- TEIXEIRA, J. (2001). *A verbalização do espaço: modelos mentais de frente/trás*. Braga: Universidade do Minho. Centro de Estudos Humanísticos.
- TURNER, M. (1991). *Reading Minds. The Study of English in the Age of Cognitive Science*. Princeton, N J: Princeton University Press.
- WALKER, L. O. (1994). *Strategies for Theory Construction in Nursing*. Prentice Hall.
- WIERZBICKA, A. (1985). *Lexicography and conceptual analyses*. Ann Arbor, MI: K.

[Recebido em 25 de janeiro de 2016 e aceite para publicação em 6 de abril de 2016]