# Semantic Specifics of Bulgarian Verbal Computer Terms

**Maria A. Todorova**

Institute For Bulgarian Language, Bulgarian Academy of Sciences,
Sofia, Bulgaria
`maria@dcl.bas.bg`

## Abstract

This paper represents a description of Bulgarian verbal computer terms with a view to the specifics of their translation in English. The study employs a subset of 100 verbs extracted from the Bulgarian WordNet (BulNet) and from the internet. The analysis of their syntactic and semantic structure is a part of a study of the general lexis of Bulgarian. The aim of the paper is to (1) identify some problem areas of the description and translation of general lexis verbs, (2) offer an approach to the semantic description of metaphor-based terms from the perspective of Frame Semantics; (3) raise questions about the definition of general lexis with respect to Bulgarian and across languages.

## 1 Introduction

This paper aims at a conceptual description of high-frequency Bulgarian verbs from the domain of computer terminology as compared with English, with some implications about the conceptual description of the relevant verbs. The goal of the analysis is to contribute both to the enrichment of the Bulgarian WordNet with Conceptual frames (Koeva, 2020) and to the enlargement of the Bulgarian FrameNet, and hence — to the creation of a linked semantic and syntactic resource.

WordNet and FrameNet are large lexical resources that provide semantic information about verb classes. WordNet (Fellbaum, 1999) represents a multilingual conceptual network of synonym sets (synsets) linked by means of semantic relations such as hypernymy, antonymy, etc. FrameNet (Baker et al., 1998) represents the semantics of lexemes by means of schematic representations (frames) describing objects, situations, or events and their components (frame elements) in the apparatus of Frame Semantics.

The hypothesis adopted in this work is that some of the term-specific words which are encountered in everyday language use belong to a semantic field (Clark, 1993) recognised as part of the general lexis.

The proposed analyses are based on the general verb lexis of Bulgarian selected for the purposes of the theoretical semantic description and typology of verb predicates belonging to the basic conceptual apparatus of the language[1] (Leseva et al., 2021; Todorova et al., 2022; Todorova, 2023). The paper also discusses some problems of domain-specific semantic representation in terms of semantic fields (Clark, 1993), abstract meaning representation of metaphor-based terms and their translation into English using WordNet as a bilingual Bulgarian - English dictionary.

The study would help to (1) identify some problem areas of the description and translation of general lexis verbs, (2) offer an approach to the semantic description of metaphor-based terms from the perspective of Frame Semantics; (3) raise questions about the definition of general lexis with respect to both Bulgarian and other languages.

The remainder of the paper is organised as follows. Section 2 describes the data used in the study – a set of computer verb terms excerpted from WordNet, a selection of manually collected verbs from the internet and a set of semantic frames. Section 3 presents the semantic features of the verb set and a comparison of the semantic descriptions of computer term verbs and their literal, non-terminological counterparts. Section 4 discusses the semantic features of the combinations of general lexis verbs and noun computer terms resulting in verb-headed computer terms. Section 5 sums up the observations on the results and suggests directions for future work.

---

[1]The selection and evaluation of the verbs that form the set of the general lexis of Bulgarian has been undertaken by the team of linguists at the Department of Computational Linguistics of the Institute for Bulgarian Language at the Bulgarian Academy of Sciences.

## 2 General Lexis and Computer Terms

It is considered that the volume of information that humanity creates and manages doubles every ten years, while the information in the area of telecommunications doubles every year. The mass penetration of telecommunications in all areas of human activity, including households has led to the demand of relatively good telecommunication literacy skills in anyone, even children. The prevalence of technology in daily life and education, along with media influence, and the overall social prestige of technological adoption results in the integration of computer terms into everyday vocabulary. Many computer terms have become so widely-used that they have entered the general vocabulary of languages. These terms are often used by people who may not have deep technical knowledge but are familiar with basic technology concepts. Words like *browse* or *click* have become as common as *iron* or *play*.

The study of verbs in the field of terminology with respect to Bulgarian has been focused mainly on the creation of new lexis (Blagoeva, 2007; Kostova, 2015) or on aspects of the metaphorical meaning of such terms (Kirova, 2018). Previous research in this field has approached the semantic description of terminological units from the perspective of Frame-based Terminology (Faber, 2012, 2015). However, this theory has not been implemented for the domain of telecommunications or Bulgarian terminology.

## 3 The Data Analysed

The computer terms were extracted from a bilingual terminological database representing 729 terms from the Computer and Technology domain in English and their translation equivalents in Bulgarian. The database[2] was created within the framework of the European Language Resource Coordination (ELRC) Connecting Europe Facility – Automated Translation (CEF.AT), Actions SMART 2014/1074 and SMART 2015/1091. The terminological resource has been compiled by combining the relevant entries in the Bulgarian WordNet – BulNet (Koeva, 2010) and entries in some other terminological monolingual dictionaries and domain-specific corpora. 319 of the terms were selected as candidates for the set of everyday lexis based on (1) the inclusion of their verb head as part of a

previously selected set of general lexis verb literals in BulNet; (2) the availability of a term in the Dictionary of neologisms[3]; and (3) the frequency of the verb in the Bulgarian National Corpus (Koeva et al., 2012). As a result, 200 nouns, 110 verbs and 9 adverbs, representing the domain of information and technology in everyday life, were extracted. As the database represents the use of the selected computer terms since 2014 and their actual use nowadays, I assume that they have been firmly established in the language as part of a stable semantic field within the general lexis set.

The descriptions of lexical semantics of verb terms are based on the combined information from the following language resources: WordNet (Fellbaum, 1999), FrameNet (Ruppenhofer et al., 2016) and the mapping of FrameNet frames and WordNet synsets (Stoyanova and Leseva, 2020).

## 4 Semantic Features

### 4.1 Semantic grouping based on semantic primitives

In order to characterise the semantic field of common computer terminology, I use the semantic domains of the verbs extracted from WordNet. These domains are grouped in 15 lexicographer's files in WordNet (Miller et al., 1990). The selected verbs are mainly from the semantic domains of verbs of contact, verbs of communication, verbs of motion, and verbs of possession.

- **Verbs of communication** refer to actions involving the exchange, transmission, and interaction of information and data between users, devices, or systems: *izprashtam* (send)[4]; *poluchavam* (receive); *prenasyam* (transfer); *spodelyam* (share); *svarzvam se* (connect); *otgovaryam* (reply).

- **Verbs of motion** describe actions related to the management, transfer, and manipulation of data and files: *kopiram* (copy); *shtrakvam klikvam* (click); *premestvam* (move); *iztrivam* (delete); *postavyam* (paste); *pridarpvam* (drag); *puskam* (drop); *otvaryam* (open); *zatvaryam* (close).

- **Verbs of possession** relate to actions involving ownership, control, and access to digital

resources and data: *ogranichavam* (restrict); *zaklyuchvam* (lock); *spodelyam* (share); *zapazvam* (save).

- **Verbs of physical contact** refer to actions involving the manipulation, operation, and interaction with hardware or devices: *natiskam* (press); *plazgam* (swipe); *razlistvam* (scroll); *vmakvam* (insert); *prikachvam* (attach); *otpechatvam* (print).

The general semantics of these groups corresponds to Conceptual frames. Most of the selected verbs have non-terminological verb counterparts. Their terminological meaning is based on metaphorical transfer which may lead to ambiguity or wrong translations. This is where the role of their description by means of semantic frames and the specification of the semantic frames of non-terminological verbs comes into play.

## 4.2 Semantic Frames

The description of the selected verbs by means of semantic frames and the representation of their semantic features – i.e. their frame elements and the relevant semantic restrictions – has been undertaken as part of the description of Conceptual frames in Bulgarian. Conceptual frames are abstract structures that describe particular types of situations or events, along with their participants and properties (Koeva, 2020). The semantic specialisation occurring within a semantic class may result in different configurations of frame elements across frames, including the inclusion or exclusion of elements, the narrowing down of their semantics (stricter selectional restrictions), etc.

Verbs in the domain of verbal computer terms refer to actions or processes involving interaction, communication and connection in the digital and technological context. The semantic field of computer terms involves a general conceptualisation of situations represented by a set of common core frame elements which may vary depending on the specific term and its context:

**Agent:** The entity that initiates or performs the action: a user, a device, a system, or a programme.
**Theme:** The entity or object which is affected by the action: a file, e-mail, a device, a network, or a system resource.
**Source:** The location or entity from which data originate in the cases of data transfer or movement: a computer or a server.

**Destination:** The location or entity to which data are transferred or sent to: a computer or a server.
**Medium:** A physical input device, software interface, network connection, etc.
**Result:** The outcome or consequence of the action. Depending on the particular verbs only the relevant frame elements are used to describe the concept. They are additionally specified by the sectional restrictions (represented in terms of semantic classes of nouns).

For example, the semantic frame of the verb *send* within the context of communication technologies is a specification of the FrameNet frame **Sending**[5] – and captures the core concepts related to the act of transmitting information or data from one location to another, which includes the following frame elements: an **Agent** – an entity which is a user, a programme or a device; a **Source** – the location or entity from which the data or information originates; a **Destination** – the location or entity to which the data or information is being transmitted; and a **Content** – the information, data, or message that is being sent from the **Source** to the **Destination**.

## 4.3 Verbal Computer Terms vs. Their Non-computer Counterparts

Verbal computer terms often share similarities with their non-computer counterparts in terms of argument structure. This allows the use of the WordNet hierarchy and its mapping with FrameNet to suggest the corresponding frame. After an appropriate frame is selected, further specify the relevant frame elements with a view to the domain of technology and telecommunications. In this way, the WordNet structure is used to enrich and make more specific the FrameNet hierarchy. However, there can be differences due to the technical nature of computing concepts and the specific actions they describe. Here are a few examples comparing computer-related terms with their non-computer counterparts.

For example, the verbs *svalyam (download)* vs. *izvlicham (retrieve)*, both described by the FrameNet frame **Removing**, differ mainly in the semantic restrictions imposed on the frame elements. The semantic frame of *svalyam (download)* is represented by an **Agent** – realised as the subject and defined as a user or a system performing the action, and a **Theme** – realised as the object and

---

defined as the data or files being transferred from a Source.The **Source** is a Location which is an adjunct of the verb and is defined as a server or a cloud from which the data is obtained.

Example: She downloaded the file from the Internet.

The semantic frame of *izvlicham (retrieve)* is represented by an **Agent** – realised as the subject and defined as a **Person** performing the action, and a **Theme** – realised as the object and defined as an **Item** being obtained. The **Source** is an adjunct of the verb and is defined as the location or origin of the **Item**.

Example: It took him some time to retrieve the file from the system. In addition, there are noun terms in the domain of computer technology which are presented in the argument structure of verbs belonging to the general lexis which do not have specialised meaning on their own. The combination of the non-terminological verb and the terminological noun results has terminological meaning itself. For example in the sentence, I'm using a mouse and headphones – the verb does not belong to the field of technology, only its arguments are computer terms. This shows specialised argument selection in the thematic domain under study and specialised use of general lexis verbs, which may pose problems to their translation via computer-aided tools.

### 4.4 Translation and Metaphors

As metaphors involve the use of a word or a phrase to describe an object or concept by comparing it to another object or concept, they play a role in enriching and expanding the general lexis of a language. Metaphors allow speakers to express abstract or complex ideas in terms of familiar and concrete concepts. Many computer term metaphors have become so deeply embedded in the language that they have lost their figurative meanings in the consciousness of speakers. People use them without recognising their metaphorical origin. In the context of digital technology, metaphors have significantly impacted general lexis. Many verbal computer terms such as *send, paste, open, close* etc. have been losing the transparency relation with their non-figurative counterparts. This is especially true for noun computer terms such as *window, virus, path, net, menu*, etc. Metaphors in the context of technology help bridge the gap between unfamiliar concepts and everyday experiences, making the terminology more accessible and relatable.

### 4.5 Translation and Adaptation

Though the set of computer terms studied in this paper belong to the specialised semantic field of computer technology, they represent a borderline case of terminological/ non-terminological lexis. In addition, they do not represent word-to-word correspondences, and a verbatim rendition in another language may lead to wrong word-to-word translation.

Most of the entries in the set of selected computer term verbs are common domestic words corresponding to the relevant English computer term. Such examples are represented by pairs such as *kachvam – upload*, *svalyam – download*, which are not terminological correspondences. This may lead to their non-terminological translation in English by another meaning of the Bulgarian verb, e.g. *kachvam – climb*, s*svalyam – put down*.

Some computer terms – borrowed foreign words have been translated or adapted to fit the linguistic patterns of the Bulgarian language(Blagoeva, 2007). This adaptation makes the terms more accessible to a wider audience with the help of domestic word-formation means employed in previously adopted loan words. For example, the English term *click* has yielded the verb *klikvam*. The word has another orthographic correspondence in Bulgarian, meaning a distance in army terminology. This raises another ambiguity problem for translation technologies.

There are also some examples of literalisation of English metaphorical terms, mainly nouns and adjectives, through their translation in Bulgarian, consider the English term *free* (software) and *bezplaten* (i.e. 'non-paid') softuer. The classes of metaphorised common lexis verbs are described in more detail by Kirova (Kirova, 2018). Determining the appropriate translation strategies of such verbs, is necessary for eliminating ambiguities and different interpretations in computer-assisted text analysis.

## 5 Discussion

One of the main points of discussions in this paper is whether computer terms used in the everyday language contribute to the general lexis. These terms have become so integrated into modern communication that they are often used in everyday conversations, even by individuals who might not consider themselves tech-savvy. As technology continues

to play a significant role in our lives, these terms are likely to become even more ingrained in the common vocabulary.

Another interesting problem rooting from the question of the place of abstract words in the general lexis set is whether metaphor-based terms are part of it.

# 6 Conclusions

This study presents a dataset of Bulgarian – English verbal computer terms from a semantic point of view which takes into account their role as a part of the relevant semantic field in the general lexis of Bulgarian. They demonstrate the integration of computer terms into everyday language and their translation and the dynamic relationship between language and technology as a result of the evolving nature of human communication. As Bulgarian is a less-resourced language, the dataset might contribute to the implementation of domain-specific tasks in computer-assisted human translation from and to this language.

Computer terms of metaphorical origin enrich the general lexis of a language and contribute to the language's evolution and the way we understand and communicate complex ideas, including those related to digital technology. The semantic description of words specific to a particular semantic field, in this case computer terms, is a specification of the semantic description of their non-terminological, non-metaphorical counterparts.

As the proposed analysis is based on interlinked multilingual language resources (WordNet and FrameNet), the observations may also be useful for other languages and may contribute to the implementation of NLP applications aimed at automatic semantic analysis, word sense disambiguation, language understanding and generation, machine translation, etc.

## Acknowledgments

# References

C. F. Baker, Ch.J. Fillmore, and John B. Lowe. 1998. The berkeley framenet project. *COLINGACL '98: Proceedings of the Conference. Montreal, Canada*, pages 86–90.

D. Blagoeva. 2007. Neologizmite v savremennia balgarski ezik. *Elektronno spisanie LiterNet*, 2.

Eve V. Clark. 1993. *The Lexicon in Acquisition.* Cambridge University Press.

P. Faber. 2012. A cognitive linguistics view of terminology andspecialized language.

P. Faber. 2015. Frames as a framework for terminology. *Kockaert, Hendrik J. and Frieda Steurs (eds.), Handbook of Terminology*, pages 14–33.

C. Fellbaum. 1999. The organization of verbs and verb concepts in a semantic net. *Text, Speech and Language Technology*, 6:278–301.

L. Kirova. 2018. Kompyutarna leksika, poluchena chrez metaforichen prenos na znachenieto na obshtoupotrebimi dumi. *Elektronno spisanie Liternet*.

S. Koeva. 2010. Bulgarian wordnet - current state, applications and prospects. *Bulgarian-American Dialogues*, pages 120–132.

S. Koeva. 2020. Towards a semantic network enriched with a variety of semantic relations. *Semantic Relations and Conceptual Frames, Koeva, S. (ed.)*, pages 7–20.

S. Koeva, I. Stoyanova, S. Leseva, T. Dimitrova, R. Dekova, and E. Tarpomanova. 2012. The Bulgarian National Corpus: Theory and practice in corpus design. *Journal of Language Modelling*, pages 65–110.

N. Kostova. 2015. Verb neologisms and respective names for actions in bulgarian language. *Balgarski ezik*, 62:48–57.

S. Leseva, I. Stoyanova, M. Todorova, and H. Kukova. 2021. Putting pieces together: Predicate-argument relations and selectional preferences. *Koeva, S. (ed.) Towards a Semantic Network Enriched with a Variety of Semantic Relations*.

G. A. Miller, R. Beckwith, D. Gross C. Fellbaum, and K. Miller. 1990. Introduction to wordnet: an on-line lexical database. *International journal of lexicography*, 3(4):235–244.

J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, and J. Scheffczyk. 2016. *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California.

I. Stoyanova and S. Leseva. 2020. Beyond lexical and semantic resources: Linking WordNet with FrameNet and enhancing synsets with conceptual frames. *Koeva, S. (ed.) Towards a Semantic Network Enriched with a Variety of Semantic Relations*, pages 21–48.

M. A. Todorova. 2023. Semantic annotation of verbs of contact in bulgarian. *Workshop on Interoperable Semantic Annotation (ISA-19)*, pages 11–17.

M. A. Todorova, T. Dimitrova, and V. Stefanova. 2022. Research on the basic verbal vocabulary in Bulgarian for students in the initial stage of education through online games. *Pedagogika-Pedagogy*, XCIV:896–913.