

СЕМАНТИЧНИ СТУДИИ.
СИНТАКТИЧНО И СЕМАНТИЧНО ОПИСАНИЕ
НА БЪЛГАРСКИ ГЛАГОЛИ



BULGARIAN ACADEMY OF SCIENCES
INSTITUTE FOR BULGARIAN LANGUAGE
PROF. LYUBOMIR ANDREYCHIN

**SEMANTIC STUDIES.
SYNTACTIC AND SEMANTIC
DESCRIPTION OF BULGARIAN VERBS**

Edited by Svetla Koeva

Reviewers: Dr. Atanas Atanasov, Dr. Rositsa Dekova

Sofia • 2024



**Prof. Marin Drinov Publishing House
of Bulgarian Academy of Sciences**

БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИНСТИТУТ ЗА БЪЛГАРСКИ ЕЗИК „ПРОФ. ЛЮБОМИР АНДРЕЙЧИН“

**СЕМАНТИЧНИ СТУДИИ.
СИНТАКТИЧНО И СЕМАНТИЧНО
ОПИСАНИЕ НА БЪЛГАРСКИ ГЛАГОЛИ**

Редактор: Светла Коева

Рецензенти: доц. д-р Атанас Атанасов, гл. ас. д-р Росица Декова

София • 2024



Издателство на БАН „Проф. Марин Дринов“

Сборникът с теоретични изследвания, озаглавен „Семантични студии. Синтактично и семантично описание на български глаголи“, е създаден по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания“, Договор № КП-06-Н 50/1 от 30.11.2020 г.

The volume “Semantic Studies: Syntactic and Semantic Description of Bulgarian Verbs” is being developed as part of the project “Enriching the Semantic Network WordNet with Conceptual Frames”, funded by the Bulgarian National Science Fund under the Program for Funding Fundamental Research (Grant Agreement No. KP-06-N50/1 of 2020).

Фонд „Научни изследвания“ не носи
отговорност за съдържанието
на отпечатаните студии.



ФОНД
НАУЧНИ
ИЗСЛЕДВАНИЯ

Адрес:

Институт за български език „Проф. Любомир Андрейчин“, Българска академия на науките

бул. „Шипченски проход“ 52, блок 17

София 1113, България

+359 2/872 23 02

Авторските права на статиите остават собственост на авторите. Публикациите в сборника се разпространяват под лиценза

Creative Commons Attribution 4.0. International Licence (CC BY 4.0): <http://creativecommons.org/licenses/by/4.0>

Editorial address:

Institute for Bulgarian Language Prof. Lyubomir Andreychin, Bulgarian Academy of Sciences

52 Shipchenski Prohod Blvd., bl. 17

Sofia 1113, Bulgaria

+359 2/872 23 02

Copyright of each paper stays with the respective authors. The works in the volume are licensed under a Creative Commons Attribution 4.0. International Licence (CC BY 4.0): <http://creativecommons.org/licenses/by/4.0>

© Институт за български език „Проф. Любомир Андрейчин“, 2024

© Издателство на БАН „Проф. Марин Дринов“, 2024

ISBN 78-619-245-448-7

ISBN online 78-619-245-448-7

СЪДЪРЖАНИЕ

Светла Коева *Семантични студии. Синтактично и семантично описание на български глаголи. Предговор* | 7

Svetla Koeva *Extended ontology of noun semantic classes* | 12

Валентина Стефанова *Част на тялото като семантична характеристика* | 46

Svetlozara Leseva, Ivelina Stoyanova *Verbs of communication: Validation of semantic frames in corpus data* | 74

Цветана Димитрова *За природата на притежателите* | 117

Мария А. Тодорова *За семантичния клас глаголи за контакт чрез преместване* | 134

Христина Кукова *Основни фреймове за перцепция и комплексно представяне на семантико-синтактична информация* | 154

TABLE OF CONTENTS

| | |
|--|-----|
| Svetla Koeva <i>Semantic studies. Syntactic and semantic description of Bulgarian verbs. Preface</i> | 7 |
| Svetla Koeva <i>Extended ontology of noun semantic classes</i> | 12 |
| Valentina Stefanova <i>Body part as a semantic feature</i> | 46 |
| Svetlozara Leseva, Ivelina Stoyanova <i>Verbs of communication: validation of semantic frames in corpus data</i> | 74 |
| Tsvetana Dimitrova <i>On the nature of Possessors</i> | 117 |
| Maria A. Todorova <i>On the semantic class of verbs of contact and displacement</i> | 134 |
| Hristina Kukova <i>Main perception frames and complex presentation of semantic and syntactic information</i> | 154 |

СЕМАНТИЧНИ СТУДИИ. СИНТАКТИЧНО И СЕМАНТИЧНО ОПИСАНИЕ НА БЪЛГАРСКИ ГЛАГОЛИ

ПРЕДГОВОР

Светла Коева

Институт за български език „Проф. Любомир Андрейчин“,
Българска академия на науките

svetla@dcl.bas.bg

Сборникът с теоретични изследвания, озаглавен „Семантични студии. Синтактично и семантично описание на български глаголи“, е създаден по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, Договор № КП-06-Н50/1 от 2020 г. В рамките на проекта се предлага интеграция на концептуална и лексикална информация с информация за синтактичните структури, в които участват лексикалните единици. По този начин резултатът е по-пълно (като обем на данните) и по-богато (от гледна точка на включената информация) описание, каквото до момента не съществува за български (а доколкото ни е известно, в такъв обем – и за други езици).

Включените в сборника изследвания са насочени към описание и типологизация на класове български глаголи, както и на съчетаемостта им с класове съществителни. Семантичното описание се основава на **система от абстрактни семантични (и концептуални) фреймове**, представящи обобщено семантичната структура на човешката представа за ситуациите от действителността, описвани от български глаголи. По дефиниция семантичните фреймове от Фреймнет са схематично представяне на знанието на говорещия за ситуациите или състоянията, които са в основата на значението на лексикалните единици (Fillmore 2007: 130). Фреймовите елементи са компоненти на семантичния фрейм и означават участниците, пропозициите, фазите и частите

* Светла Коева. Семантични студии. Синтактично и семантично описание на български глаголи. Предговор. – В: Св. Коева (ред.) *Семантични студии. Синтактично и семантично описание на български глаголи*. София: Издателство на БАН „Проф. Марин Дринов“, с. 7 – 13 <https://doi.org/10.7546/SemanticStudies2024.01>

от ситуациите (Fillmore 2007: 130). Синтактичното описание представя в пълнота (или някои от възможностите за) синтактичната реализация на фреймовите елементи от съответните семантични (и концептуални) фреймове.

Фундаменталният характер на изследванията се предполага от дефинирането на понятието *концептуален фрейм*: абстрактна структура, която може да съответства в пълнота или частично на семантичен фрейм от Фреймнет или да е създадена по образец на семантичните фреймове за описание на специфични за българския език лексикални единици. В допълнение, в концептуалните фреймове елементите на фрейма са асоциирани със синонимни множества на съществителни имена, които доминират над подходящо за съчетаване с глагола множество от думи (Коева 2020: 7).

Абстрактното семантично моделиране води до теоретични обобщения за онтологичното описание на семантичните класове глаголи, за системата от семантични (и концептуални) фреймове, към които принадлежат, както и за комплексното представяне на семантичната информация.

Част от студиите, включени в сборника, са посветени на теоретични обобщения за системата от семантични класове на съществителните, подходящи за комплексното описание на допустимите съчетания от глаголи и съществителни в различни изреченски позиции: студията на Светла Коева и студията на Валентина Стефанова.

В резултат на корпусни изследвания и лингвистично моделиране е представена *Разширена онтология на семантичните класове на съществителните в Уърднет* (доколкото е възможно, езиково независима), свързана с възможностите за тяхното съчетаване с различни типове предикати (студията на Светла Коева *Extended ontology of noun semantic classes*). Елементите от *Разширената онтология на семантичните класове на съществителните* са съотнесени с подходящи синонимни множества от Уърднет (Miller et al. 1993), семантичните типове от Корпусния анализ на изреченски модели (Hanks 2012), семантичните типове от Фреймнет (Fillmore, Baker 2010) и селективните предпочитания от Върбатлас (Di Fabio et al. 2019). Това прави възможно типологичното описание на предикатите и открояването на езиково специфичните и универсалните особености на семантично и синтактично равнище. Представен е моделът за изграждане на Българския фреймнет, в който са възприети две абстрактни структури: суперфрейм и концептуален фрейм. Суперфреймовете се конструират от семантичните фреймове на Фреймнет посредством отстраняване на информацията, специфична за английски. Концептуалните фреймове, от своя страна, се създават по модела на семантичните фреймове, но включват еквивалентно описание за всички лексикални единици, които ги предизвикват.

Студията на Валентина Стефанова *Част на тялото като семантична характеристика* се фокусира върху изследване на фреймовия елемент **Част на тялото** и семантичните фреймове, към които принадлежи. Определени са глаголните синонимни множества от Българския уърднет, чиито лите-

рали предизвикват представата за семантични фреймове с ядрен фреймов елемент **Част на тялото**. Изследването предлага групиране на съществителните имена, които се съчетават с глаголите при реализация на изследвания фреймов елемент, посредством приписване на семантични типове от Корпусния анализ на изреченските модели към синонимните множества на съществителните от Уърднет. Прави се обобщението, че в повечето случаи няма ограничения за лексикалното проектиране на фреймовия елемент **Част на тялото** посредством думите, които обозначават части на тялото. След внимателен преглед се предлага по-конкретна класификация на съществителните имена, пренадлежащи към различни семантични типове, в зависимост от конкретния семантичен фрейм, който специфицира употребата в определен контекст.

Друга част от студиите в сборника (на Светлозара Лесева и Ивелина Стоянова и на Цветана Димитрова) предлагат теоретични обобщения за системата от семантични (и концептуални) фреймове. Теоретично значение има съпоставката на концептуализацията на различни типове предикати, която корелира с граматичната реализация на съответните фреймови елементи и семантичната им съчетаемост.

В студията *Verbs of communication: Validation of semantic frames in corpus data* от Светлозара Лесева и Ивелина Стоянова се изследват глаголите за комуникация с фокус върху няколко основни семантични фрейма за комуникация във Фреймнет. Анализът е съсредоточен върху семантичното представяне и валентните модели, показващи комбинаториката между фреймовите елементи и синтактичната им реализация, за английски и български. Въз основа на корпусни данни, извлечени от семантично анотираните корпуси за английски и за български, се подлагат на анализ семантичните свойства и синтактичната реализация на групи глаголи, принадлежащи към няколко представителни семантични фрейма за комуникация. Наред с това се извеждат и основни прилики и разлики при реализацията както между български глаголи, така и между тях и съответствията им в английски.

Студията *За природата на притежателите* от Цветана Димитрова предлага преглед на начините за изразяване на **притежател** чрез съществителни имена от определени семантични класове, които отговарят на изискванията за реализация на съответните фреймови елементи от семантичните фреймове на Фреймнет. Разглеждат се реализациите на фреймови елементи, които изразяват различни притежатели, при основни глаголи за трансфер на притежание. По-конкретно, център на внимание са фреймовият елемент, който е обект на трансфера на притежание, и двата фреймови елемента, между които протича трансферът, както и някои периферни фреймови елементи. Въз основа на анализ на корпусни примери, извлечени със СкетчЕнджин, се аргументира дефинирането на множеството от семантични класове съществителни имена, подходящи за реализацията на разглежданите фреймови елементи.

Друга част от студиите са посветени на теоретични обобщения за комплексното представяне на семантична информация (на Мария Тодорова и на Христина Кукова).

Студията на Мария А. Тодорова *За семантичния клас глаголи за контакт чрез преместване* предлага описание на глаголите за **контакт**, комбиниран с преместване, които са включени в Уърднет и принадлежат към общоупотребимата лексика в български. По-конкретно, изследването се фокусира върху глаголите за прикрепване, отделяне, поставяне, премахване, пълнене и изпразване. Предложено е семантико-синтактично описание, което обхваща фреймовите елементи от разгледаните семантични фреймове от Фреймнет и съответстващите им синтактични реализации. Обръща се внимание на семантичните ограничения при възможната лексикална реализация на фреймовите елементи. Представената информация се основава на ръчна проверка, редакция и допълване на полуавтоматичното съотнасяне на Уърднет и Фреймнет.

Студията на Христина Кукова *Основни фреймове за перцепция и комплексно представяне на семантико-синтактична информация* е посветена на перцептивните глаголи и концептуалното им описание в българския език. Като отправна точка за проучването служи разнообразието от конструкции и синтактични реализации, които представителите на този клас предикати проектират върху изреченското си обкръжение. Разгледани са седем семантични фрейма от Фреймнет, чиито лексикални единици демонстрират висока честота и голямо разнообразие. Описанието включва граматични особености на глаголите, семантични и синтактични ограничения, които глаголните лексикални единици налагат при реализацията на фреймовите елементи, съотнасяне на фреймовите елементи с определено синонимно множество или онтологичен клас от съществителни имена от Уърднет.

Изследванията по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“ ще допринесат за изясняване на характера на семантичната структура и реализацията на фреймовите елементи на синтактично равнище с оглед на съвременното състояние на българския език.

Не на последно място концептуалното описание има значение за моделирането на семантичните отношения между лексикални единици и фреймови елементи и може да допринесе **при решаването на проблеми в областта на компютърната обработка на естествения език, които са актуални в световен мащаб, като:** автоматично маркиране на семантични роли, автоматичен семантичен и синтактичен анализ, извличане на информация, автоматично отговаряне на въпроси, автоматичен превод и др.

Цитирана литература / References

Di Fabio et al. 2019: Di Fabio, Andrea, Simone Conia, and Roberto Navigli. VerbAtlas: a novel large-scale verbal semantic resource and its application to semantic role labeling. – In: *Proceedings of the 2019 Conference on Empirical Methods in Natural*

- Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China: Association for Computational Linguistics, pp. 627 – 637.
- Fillmore 2007: Fillmore, Charles J. Valency issues in FrameNet. – In: Herbst, T., K. Gutz-Votteler (Eds.). *Valency: Theoretical, descriptive and cognitive issues*. Berlin: Mouton de Gruyter, pp. 129 – 160.
- Fillmore, Baker 2010: Fillmore, Charles J. and Collin Baker. A frames approach to semantic analysis. – In: Bernd Heine and Heiko Narrog (Eds.) *The Oxford handbook of linguistic analysis*, Oxford: Oxford University Press, pp. 313 – 340.
- Hanks 2012: Hanks, Patrick. How people use words to make meanings: Semantic types meet valencies. – In: A. Boulton and J. Thomas (Eds.) *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno: Masaryk University Press, pp. 54 – 70.
- Koeva 2020: Koeva, Svetla. Semantic Relations and Conceptual Frames. – In: Koeva, S. (Ed.). *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Institute for Bulgarian Language Prof. Lyubomir Andreychin, pp. 7 – 20.
- Miller et al. 1993: Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. *Introduction to WordNet: an On-line Lexical Database. Five Papers on WordNet*. Princeton, NJ: Princeton University.

Svetla Koeva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52 Shipchenski prohod Blvd., bl. 17
Sofia 1113, Bulgaria
<https://orcid.org/0000-0001-5947-8736>

EXTENDED ONTOLOGY OF NOUN SEMANTIC CLASSES

Svetla Koeva

Institute for Bulgarian Language „Prof. Lyubomir Andreychin“ –
Bulgarian Academy of Sciences

svetla@dcl.bas.bg

Abstract. The primary aim of this study is to propose an effective approach for characterising the frame elements of conceptual frames within the Bulgarian FrameNet using noun classes. This approach will facilitate the prediction of comprehensive and semantically valid combinations between verbal lexical units that evoke conceptual frames and appropriate nouns. The study provides a concise overview of the semantic classifications of nouns in WordNet, Corpus Pattern Analysis, FrameNet, and VerbAtlas, emphasising their significance in predicting verb-noun compatibility. The structure of the Bulgarian FrameNet (BulFrame) is presented, containing components adapted from FrameNet for Bulgarian along with a substantial amount of lexical, morphological, syntactic, and semantic information specific to the Bulgarian language. One distinctive feature of the Bulgarian FrameNet is the specification of noun classes, indicating appropriate nouns for the lexical realisation of frame elements. By aligning synonym sets from WordNet with the semantic types of Corpus Pattern Analysis and FrameNet, as well as with the selective preferences of VerbAtlas, the foundational structure of the Extended Ontology of Noun Semantic Classes is established. The Ontology concepts are linked, not exclusively, to synonym sets from WordNet and thereby to sets of nouns suitable for combination with verbal lexical units that evoke the corresponding conceptual frames. The contribution of this development lies in detailing the steps for selecting semantic classes (concepts) and constructing the structure of the Extended Ontology of Noun Semantic Classes.

Keywords: *semantic class, semantic type, selectional preferences, conceptual frame, frame element*

* Svetla Koeva. Extended ontology of noun semantic classes. – In: Koeva, Sv. (Ed.) *Semantic studies. Syntactic and semantic description of Bulgarian verbs*. Sofia: BAS Press „Prof. Marin Drinov“, pp. 12 – 45. <https://doi.org/10.7546/SemanticStudies2024.02>

1. Introduction

There are different (ontological) representations of word classes, each with its own set of concepts and degree of complexity. For example, nouns and verbs in WordNet are categorised into semantic classes (Miller 1990/1993: 16 – 17), nouns in Corpus Pattern Analysis are classified into semantic types (Hanks 2012: 57 – 17), frame elements in FrameNet are often specified according to the semantic types of nouns (Ruppenhofer et al. 2016: 86 – 87), and semantic roles in VerbAtlas are supplied with selectional preferences (Di Fabio et al., 2019: 630).

The objective of this study is to propose an effective approach for characterising the conceptual frame elements in the Bulgarian FrameNet using noun classes. This method is designed to streamline the prediction of comprehensive and semantically sound pairings between verbal lexical units, which evoke conceptual frames, and appropriate nouns.

The study offers a succinct overview of the semantic classifications of nouns in WordNet, Corpus Pattern Analysis, FrameNet, and VerbAtlas, underscoring their pivotal role in anticipating verb-noun compatibility. Furthermore, as part of a broader investigation, it sheds light on the inadequacies of existing noun classifications, which are often unsuitable for effectively illustrating syntagmatic verb-noun combinations.

The structure of the Bulgarian FrameNet (BulFrame) is introduced, comprising components tailored within FrameNet, complemented by a significant array of lexical, morphological, syntactic, and semantic information unique to the Bulgarian language. A notable aspect of the Bulgarian FrameNet is its delineation of noun classes, providing predictions on the lexical realisation of frame elements.

Drawing from existing language resources, the study offers a comprehensive approach that combines appropriate ontological representations of noun classes and extends them with additional classifications. By aligning synonym sets from WordNet with the semantic types of Corpus Pattern Analysis and FrameNet, as well as with the selective preferences of VerbAtlas, the foundational structure of the Extended Ontology of Noun Semantic Classes is established. The Ontology concepts are linked, not exclusively, to synonym sets from WordNet and thereby to sets of nouns suitable for combination with verbal lexical units that evoke the corresponding conceptual frames. The contribution of this development lies in detailing the steps for selecting semantic classes (concepts) and constructing the structure of the Extended Ontology of Noun Semantic Classes.

The paper is structured as follows: the subsequent four sections offer an overview of the semantic classifications of nouns in WordNet, Corpus Pattern Analysis, FrameNet, and Bulgarian FrameNet, focusing on their relevance to verb-noun compatibility. Section six provides a brief discussion of appropriate resources presenting groupings of nouns that can co-occur with specific verbs and outlines our approach to identifying noun classes with appropriate candidates for particular verb-noun compatibilities. The initial development of the Extended

Ontology of Semantic Classes of Nouns entails mapping relevant ontologies to the WordNet noun hierarchy, facilitating the identification of concepts indicating appropriate noun fillers for specific frame elements.

The contribution of this investigation lies in the detailed and systematic steps employed to establish a suitable set of semantic classes organised within the Extended Ontology of Semantic Classes of Nouns. Additionally, the mapping with Corpus Pattern Analysis semantic types, VerbAtlas selectional preferences, and FrameNet semantic types highlights the importance of high-rated concepts. The alignment of these resources emphasises the relevance of certain semantic classes. Furthermore, it illustrates how the Ontology of Semantic Classes of Nouns remains open to augmentation.

2. FrameNet, Semantic Frames and Semantic types

FrameNet is based on the theory of Frame Semantics (Fillmore 1976, 1982; Fillmore, Baker 2010). The core concept of Frame Semantics revolves around describing word meanings in relation to semantic frames, which are seen as schematic representations of conceptual structures within a speech community (Fillmore et al. 2003: 235). Semantic frames, more succinctly, represent speakers' understanding of the underlying situations or states of affairs shaping the meanings of lexical items (Fillmore 2007: 130). When a lexical unit evokes a frame, it invokes the associated conceptual structure, while a valency description of a specific lexical unit outlines how the semantic valences are expressed within sentences constructed around the frame-bearing unit (Fillmore 2007: 131).

FrameNet functions as a repository of semantic frames and valency patterns associated with lexical units, offering conceptual-semantic and syntactic descriptions derived from annotated examples. Frame Semantics aligns target lexical units with linguistic and conceptual details. Linguistic facets encompass definitions, semantic classes, and syntactic patterns for frame elements, while conceptual aspects involve situational descriptions, participant delineations, and inter-frame relations (Sikos, Pado 2018: 2).

2.1 FrameNet semantic frames

In FrameNet, the **semantic frame** comprises several key components: the frame name, an informal **definition** of the represented situation, an optional specification for the semantic type, a set of associated **frame elements** (including core, core-unexpressed, and non-core elements like peripheral and extrathematic), specifications for **relations between frame elements** and **frame-to-frame relations** if applicable, and the **lexical units** evoking the frame. Information on frame elements includes their names, informal definitions, optional semantic types, and illustrative examples. Details for lexical units encompass definitions, optional

semantic types, examples supplied with **annotations covering frame elements, grammatical categories, and functions.**

While certain **frame element names** may align with semantic roles, these names primarily function as mnemonic aids (Fillmore et al. 2003: 237). **Frame element definitions** are statements that convey a frame element's semantics with regard to the target lexical unit (and, in some cases, to other frame elements as well).

Two criteria are employed in formulating semantic frames (Ruppenhofer et al. 2016: 11 – 17): a checklist of features and additional principles such as paraphrases and alternative answers to a question. The checklist of features encompasses:

- Consistent number and type of frame elements across all lexical units.
- Uniform set of stages and transitions (subevents) shared by lexical units. For example, while *decapitate* implies death, *shoot* merely implies firing and hitting.
- Uniform participant perspective, such as the buyer's or seller's viewpoint.
- Consistent interrelations between frame elements for all lexical units.
- Similar presuppositions, expectations, and concomitants of the target lexical units.
- Similar basic denotation among lexical units.
- Consistent pre-specifications given to frame elements by frame-evoking lexical units. For example, verbs like *crowd*, *flock*, *pour*, and *stream* are part of the **Mass motion** frame but not the **Self motion** frame, as they imply movement by a mass theme.

The development of frames is also based on the paraphrasability (or near-paraphrasability) of lexical units: whether one lexical unit can be more or less successfully substituted for another while evoking the same frame and the same configuration of frame elements ((Ruppenhofer et al. 2016: 15 – 17). One and the same semantic frame can be evoked by synonyms, near synonyms, antonyms, derivationally related lexical units, hypernyms, or hyponyms (Koeva 2021: 183).

In FrameNet, frame elements are categorised based on their centrality within a given frame, delineating three levels: core, peripheral, and extrathematic. A **core frame element** is indispensable to the frame's central meaning (Fillmore 2007: 133) and embodies a conceptually essential component that distinguishes the frame from others (Ruppenhofer et al. 2016: 23). **Peripheral frame elements** denote notions like Time, Place, Manner, Means, and Degree, devoid of distinctiveness among frames and applicable to any semantically appropriate frame (Ruppenhofer et al. 2016: 24). **Extrathematic frame elements**, while present, aren't conceptually inherent to the frame; they relate to other abstract frames and contextualise events against the backdrop of another event (Fillmore 2007: 133). **Core-unexpressed frame elements** function as core elements but remain unexpressed in descendants of the frame; they're absorbed by lexical units in child frames, lacking individual representation (Ruppenhofer et al. 2016: 25).

Observations indicate that frame elements may not operate independently; certain groups behave akin to sets, termed **Core Sets**, wherein the presence of any member suffices to fulfil the predicator’s semantic valence (Ruppenhofer et al. 2016: 25). For example, in motion frames, **Source**, **Path**, and **Goal** core frame elements constitute a **Core Set**, allowing for the appearance of one or two (rarely all three) frame elements in a sentence without violating the semantic structure. The relation **Requires** is applied when the presence of one core frame element necessitates the occurrence of another core frame element. Conversely, the relation **Excludes** emerges when one frame element from a set of conceptually related elements is present, precluding the presence of any other element from that group (Ruppenhofer et al. 2016: 26). For instance, in the **Attaching** frame, the frame elements **Goal** and **Item** are mutually required to complement each other, while excluding the presence of the frame element **Items**.

As it has been pointed out, the semantic and syntactic descriptions in FrameNet differ from other lexical resources in several ways (Fillmore 2007: 129), including: (1) relying on corpus evidence; (2) basing the semantic layer of valency on an understanding of the cognitive frames that motivate and underlie the meanings of each lexical unit; (3) recognising various kinds of discrepancies between lexical units on the semantic (functional) level and patterns of syntactic form; and (4) providing the means of assigning partial interpretations to valents that are conceptually present but are syntactically unexpressed.

2.2 FrameNet Semantic Types

Frames, frame elements, and lexical units are categorised by ontological semantic types. On frames, the semantic type indicates that every lexical unit of the frame could be labeled with an equally or more specific type. The **Clothing** frame, for example, has the semantic type [Artefact]. As a result, all of its lexical units (boot.n, cape.n, dress.n, and so on) denote artefacts (Loenneker-Rodman, Baker 2009: 422 – 423).

Semantic types for frame elements classify the type of filler that is expected to appear as the frame element (Ruppenhofer et al. 2016: 86). Not all frame elements (and frames) come with a specified semantic type, and in general, the semantic types tend to be too broad, lacking precision in conveying actual restrictions for lexical combinations. For example, certain frame elements within the semantic frame **Experiencer focused emotion** have rather general semantic types: **Content** with the semantic type [Content]; **Event** with the semantic type [State of affairs]; **Experiencer** with the semantic type [Sentient]; **Degree** with the semantic type [Degree]; **Explanation** with the semantic type [State of affairs]; **Manner** with the semantic type [Manner]; **Time** with the semantic type [Time]. On the other hand, some frame elements, such as **Topic**, **Expressor**, **State** are not specified with a semantic type (Koeva 2021: 187).

The semantic types in FrameNet provide a high level of abstraction to express the multiple entities that can fill a frame element. A common semantic type in

FrameNet is [Sentient], which is associated with frame elements that typically represent humans or sentient entities. For example, in the semantic frame **Telling** (with a definition ‘A **Speaker** addresses an **Addressee** with a **Message**, which can be indirectly referred to as a **Topic**. Instead of (or in addition to) a **Speaker**, a **Medium** may also be mentioned.’), both frame elements **Speaker** and **Addressee** have the semantic type [Sentient], although only nouns referring to humans can be the appropriate fillers when used non-metaphorically.

Most semantic (ontological) types in FrameNet can be directly mapped to WordNet synset nodes and ontologies. They are organised hierarchically, with inheritance relations (or “is-a” relations) between parent nodes and child nodes (Ruppenhofer et al. 2016: 86). At the top level, there are five distinct types: [Attribute], [Physical entity], [Event], [Group], and [Relations]. The remaining forty ontological semantic types are subsumed under one of these five top-level types. For example, [Speed] is a type of [Attribute]; [Human] is a type of [Sentient], which is a type of [Animate being], which is a type of [Living thing], which is a type of [Physical object], which is a type of [Physical entity]; [Activity], [Accomplishment] and [Achievement] are types of [Event], which is a type of [State of Affairs]; [Organisation] is a type of [Group]; [Source], [Path] and [Goal] are types of [Locative relation] which is a type of [Relations], and so forth.

As evident, FrameNet semantic types offer a high level of abstraction and, in many instances, prove challenging to employ for predicting diverse context realisations.

3. WordNet, Sentence Frames and Semantic Classes (Primitives)

WordNet is a semantic network whose nodes host synonyms denoting different concepts and whose arcs, connecting the nodes, encode different types of relations (semantic: genus-kind, part-whole, cause-effect, etc.; morpho-semantic: agent-predicate, predicate-instrument, predicate-state, etc.; derivational; extralinguistic, i.e., membership to a thematic domain; inter-language, i.e., translation equivalents). The idea for organising the lexicon of a given language into a (lexico-)semantic network was first executed in the Princeton WordNet¹ (Miller et al. 1990/1993). Some of the fundamental ideas on which the WordNet is based encompass: a) the use of a semantic network that embraces taxonomies, meronomies, and non-hierarchical relations with clearly defined properties that allow for quick and easy automatic processing; b) a different organisation of the lexicon in comparison to traditional dictionaries where words are ordered alphabetically and the links among semantically related words (such as between sister hyponyms, between a whole and its parts, etc.) are not explicitly presented (Miller 1986).

¹ <https://wordnet.princeton.edu>

3.1. WordNet sentence frames

WordNet presents one or more sentence frames for each verb synset to clarify the basic syntactic properties of verbs – the subcategorization features of verbs by indicating the types of sentences they can occur in (Fellbaum 1990/1993: 55). These sentence frames detail semantically (and syntactically) obligatory constituents and constraints on combinability. For example, the synset {read} ‘interpret something that is written or printed’ is associated with three sentence frames:

Somebody ----s

Somebody ----s something

Somebody ----s that CLAUSE

while synset {assure, tell} ‘inform positively and with certainty and confidence’
– with two:

Somebody ----s somebody of something

Somebody ----s that CLAUSE

Apart from the number of the syntactically obligatory constituents, minimal details concerning the specific combinability restrictions – that is, whether a particular sentence frame element can be realised as a human noun or not – and minimal syntactic details – that is, whether the element is realised as a noun, prepositional phrase, or clause – are provided.

3.2. WordNet semantic classes

WordNet categorises concepts and arranges them in a hierarchical structure using a series of semantic primitives (semantic classes) for nouns and verbs (Miller 1990/1993: 16; Fellbaum 1990/1993: 41). These semantic classes serve as the basic building blocks for capturing essential semantic distinctions and categorising words based on their meanings.

Nouns are categorised into twenty-five semantic classes: {**act**; action; activity}, {**animal**; fauna}, {**artifact**}, {**attribute**; property}, {**body**; corpus}, {**cognition**; knowledge}, {**communication**}, {**event**; happening}, {**feeling**; emotion}, {**food**}, {**group**; collection}, {**location**; place}, {**motive**}, {**natural object**}, {**natural phenomenon**}, {**person**; human being}, {**plant**; flora}, {**possession**}, {**process**}, {**quantity**; amount}, {**relation**}, {**shape**}, {**state**; condition}, {**substance**}, {**time**} (Miller 1990/1993: 16). Verbs are classified into fifteen semantic classes: fourteen classes for events or actions (verbs of bodily care and functions, change, cognition, communications, competition, consumption, contact, creation, emotion, motion, perception, possession, social interaction, and weather verbs), and one class for verbs denoting states (Fellbaum 1990/1993: 57 – 61).

Nouns are categorised according to semantic classes, with a (relatively small) number of basic concepts selected and each treated as the unique starting point of a separate hierarchy. These hierarchies correspond to relatively independent semantic fields, each with its own vocabulary (Miller 1990/1993: 16). Noun-

hypernym subtrees can be linked in a hierarchical structure to the most abstract root {entity} ‘that which is perceived or known or inferred to have its own distinct existence (living or nonliving)’. There are over 550 verb subtrees in WordNet, and the same semantic class can occur in many subtrees. The verb subtrees can be artificially combined into an abstract root *eventuality*, not present in WordNet and comprising events, processes, states, changes, etc.

Although (at least for nouns) the idea is that the semantic classes in WordNet should be differentiated in order to develop multiple hierarchies organising words with different semantic classes, the WordNet structure includes nouns and verbs of different semantic classes within a single hypernymy tree. For example, for nouns, the synset {act; human action; human activity} ‘something that people do or cause to happen’ has the semantic class *noun.act*, while its hyponym {communication; communicating} ‘the activity of communicating; the activity of conveying information’ has the semantic class *noun.communication*. Among the verbs, {interact} ‘act together or towards others or with others’ has the semantic class *verb.social*, while its troponym {communicate; intercommunicate} ‘transmit thoughts or feelings’ has the semantic class *verb.communication* and the next level troponym {grimace; make a face; pull a face} ‘contort the face to indicate a certain mental or emotional state’ has the semantic class *verb.body*.

Further, the semantic class *noun.communication* applies to noun synsets, including {communication; communicating}, {language; linguistic communication}, {visual communication}, {email; e-mail; electronic mail}, {body language} and so on. The semantic class *verb.communication* specifies verb synsets such as {communicate; pass on; pass; pass along; put across}, {articulate; enunciate; vocalise}, {sign; signal; signalise}, {broadcast; air; transmit; beam; send}, etc. It is intuitively obvious that nouns and verbs referring to verbal and non-verbal communication belong to different semantic subclasses; the same applies to written communication, communication by signalling, communication by broadcasting, and so on. From this, it can be concluded that there is still much potential for additional specification of WordNet’s semantic classes in order to capture the semantic compatibility of verbs and nouns.

Not all verbs marked with the semantic class *verb.communication* can be combined with the nouns of the class *noun.communication* as objects. Based on an analysis of the superordinates (hypernyms), it has already been observed that the nodes in the WordNet hierarchy do not (always) represent semantic classes, nor do those classes occupy specific slots in the verb argument structure (Hanks, Pustejovsky 2005: 66).

WordNet has been used as an ontology (lightweight ontology) in some applications based on the fact that the hypernymy hierarchies represent subsumption between concepts (Basile 2015). In addition, there are attempts (OntoWordNet) to convert WordNet into an ontology in two phases: automatic phase, in which WordNet glosses are parsed and an approximate definition of WordNet concepts is generated, in which generic associations (A-links) are established between the concept and other concepts

that appear in its gloss; and partially automatic phase, in which the foundational top ontology DOLCE (in its DOLCE-Lite+ version with about 300 concepts) (Borgo et al. 2022) is used to interpret A-links in the form of axiomatized conceptual relations (Gangemi et al. 2003). WordNet has also been mapped several times to different ontologies: SUMO (Niles, Pease 2003); KYOTO (Laparra et al. 2012), etc.

The fact that WordNet can be mapped to pre-existing ontologies or that it can be transformed into an ontology shows that hypernymy inheritance between concepts has the character of an ontological representation, although some formal requirements for ontologies are not fulfilled. The two observations that the WordNet semantic classes can be subdivided into subclasses and that the taxonomic structure of a WordNet is close to an ontological representation lead to the view that, on the one hand, hierarchies of nouns in a WordNet are appropriate for selecting sets of nouns suitable for pairing with verbs in their valency slots, and that general semantic classes of nouns can be subdivided into appropriate subclasses appropriate for noun fillers' specification.

4. Corpus Pattern Analysis, Verb Patterns and Semantic Types

Corpus Pattern Analysis (CPA) is a technique for linking word meaning and usage (Hanks, Pustejovsky 2005: 64). The key idea is that while words have multifaceted potential to contribute to the meaning in a context, they have no specific meaning when used alone, and depending on the situation, different aspects of this meaning potential are realised. Evidence from corpuses demonstrates that the contextual patterns of word use are very regular.

4.1. CPA verb patterns

The Corpus Pattern Analysis is used to create a Pattern Dictionary of English Verbs (PDEV). A PDEV verb entry consists of a list of numbered patterns (frames) that are linked to implicatures – explanations of the meaning of the patterns (Hanks 2004: 88). For example, there are eleven registered patterns of the verb *grasp*, three of which are present here:

Human | Animal **grasp** Physical_Object

[[Human | Animal]] seizes [[Physical_Object]] and holds it firmly

Human 1 **grasp** Human 2 | Body_Part | Garment

[[Human 1]] seizes [[Body_Part | Garment]] of [[Human 2]] for some purpose

Human **grasp** at | for Physical_Object

[[Human]] attempts to seize [[Physical_Object]]

4.2. CPA semantic types

Collocation analysis enriches understanding beyond valency and captures nuanced semantic relationships between words (Hanks 2012: 58-60). By identifying statistically significant collocates, they can be organised into lexical sets sharing semantic features, such as [Human]. Each verb pattern's slot is represented by a lexical set of nouns, ranging from single words to extensive collections (Hanks 2012: 62).

The patterns are comprised of structured sentence roles, typically filled with nouns sharing aspects of their meaning, categorised as semantic types (Cinkova, Hanks 2010: 4; Hanks 2012: 66). A semantic type serves as a classification to which words can be assigned; for example, *Peter* or *the old man* are classified under the semantic type [Person] (Hanks, Pustejovsky 2005: 64). Essentially, semantic types such as [Human], [Animal], or [Part], generalise properties expressed by frequently encountered words in specific pattern positions (Hanks 2012: 57 – 59).

The initial iteration of the CPA ontology comprises a shallow ontology featuring 65 semantic types chosen for their prevalence in a manually identified selection of context patterns (Pustejovsky et al. 2004). The ontology of CPA semantic types is expandable, accommodating the addition of new types as they surface within emerging verb patterns, currently numbering 253 types. While some verb patterns demonstrate broad preferences, such as [Anything], others delineate preferences for a restricted set of words grouped into semantic types. These semantic types express the semantic preferences dictating the array of nouns and noun phrases commonly found in a given sentence position.

The verb *ask*, for example, is linked to twelve verb patterns, in most of which the semantic types are [Human], [Institution], and [Anything]:

Human 1 **ask** Human 2 QUOTE WH+

[[Human 1]] says {QUOTE} to ([[Human 2]]) in the form of a question, for example because [[Human 1]] wants to find out

Human 1 | Institution 1 **ask** Human 2 | Institution 2 question about Anything

[[Human 1 | Institution 1]] puts a {question} to ([[Human 2 | Institution 2]]) in order to find out ({about [[Anything]])}

Some verb patterns accept only a small selection of lexical units (in some cases, a word) as noun collocates, and no semantic type is defined; instead, the lexical units are listed in the verb pattern. For example, the word *permission* fits the following pattern for the verb *ask*:

Human 1 **ask** Human 2 permission to+INF

[[Human 1]] formally requests permission from [[Human 2]] to be allowed {to/INF [V]}

Semantic types, in contrast to WordNet semantic classes, stem from corpus-driven generalisations regarding noun groupings based on their collocations

with specific verbs. CPA semantic types outnumber WordNet semantic classes approximately tenfold and can be further extended. While semantic types embody cognitive concepts grounded in corpus evidence, they retain abstract status without associations with collections of nouns belonging to their respective classes. Initially, the set of 65 semantic types was linked with around 20,000 nouns (Pustejovsky et al. 2004), but the details are not available. Hence, the task of relinking remains an ongoing endeavour.

5. Bulgarian FrameNet, Conceptual Frames and Semantic Classes

The efforts towards establishing the Bulgarian FrameNet span approximately 20 years, with its origins tracing back to predecessors like the Valency Dictionary for Bulgarian and the Semantic-Syntactic Dictionary for Bulgarian (Koeva et al. 2003). Initially, the focus of resources on frame-like semantic and syntactic descriptions was exclusively on Bulgarian, without establishing connections with FrameNet.

In subsequent phases, appropriate semantic frames were manually selected, and language-independent information was extracted from these frames. This data was then enriched with Bulgarian lexical units evoking corresponding frames, along with annotated examples (Koeva et al. 2008; Koeva 2010). However, this work encountered challenges, including the lack of appropriate means to maintain correspondence with semantic frames while reconstructing them to adequately represent certain Bulgarian lexical units. Additionally, challenges arose in encoding translation equivalence between Bulgarian and English lexical units and ensuring annotation consistency regarding Bulgarian grammatical structure.

At its current stage, the Bulgarian FrameNet incorporates two abstract semantic structures: a superframe and a conceptual frame. It includes lexical units accompanied by comprehensive lexical, semantic, and grammatical information that evoke conceptual frames, along with valency patterns derived from authentic examples.

The primary aim of introducing superframes and conceptual frames is to incorporate language-specific information while ensuring alignment with relevant semantic frames. Superframes establish abstract mappings between semantic frames in FrameNet and their Bulgarian counterparts, serving as a bridge between semantic resources. Conceptual frames, linked with a specific superframe, encode relevant information for Bulgarian, which may fully or partially overlap with their English counterparts (Figure 1).

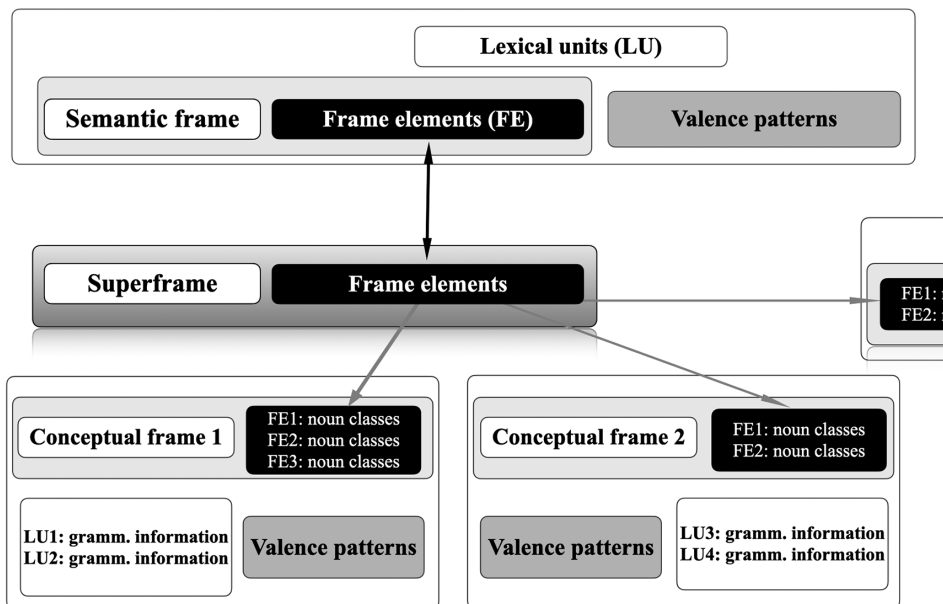


Figure 1. Organisation of data in the Bulgarian FrameNet

5.1 Superframes and conceptual frames in BulFrame

By introducing an intermediary abstract layer like superframes, the appropriate components in the Bulgarian FrameNet can align with FrameNet semantic frames while preserving specificity where needed. **Superframes** are crafted by stripping away all language-specific details for English, such as lexical units evoking the frames, their parts of speech and lexical types, and English sentences illustrating the frame and frame element definitions. Instead, they retain only non-language-specific information: semantic frames, their semantic types and definitions, frame-to-frame relations, frame elements, their semantic types and definitions, frame element relations, and administrative details like frame and frame element names.

This approach aims to establish a seamless connection with FrameNet while facilitating the identification and outlining of language-specific conceptualisations unique to Bulgarian. It also allows for the possibility of splitting one semantic frame into multiple conceptual frames, each characterised by different levels of reconstruction. An equivalence relation is established between the language-independent information in a semantic frame and its counterpart in a superframe.

Conceptual frames serve to introduce script-like descriptions relevant to Bulgarian, which may be entirely or partially analogous to the information for English or unique (in rare cases), providing conceptual descriptions specific to Bulgarian. Consequently, one superframe may be associated with one or more conceptual frames. However, there can be at most one conceptual frame whose language-independent

components are related to the superframe and, consequently, to the semantic frame through an equivalence relation. The remaining conceptual frames are linked to the superframe by partial equivalence relations, allowing for tracing to determine which components align with those in the superframe and which do not.

A **conceptual frame** can (similarly to the semantic frame) be defined as an abstract structure that describes a certain type of situation or event, along with its actors and properties (Koeva 2020: 7). The **conceptual frame** is characterised by frame elements, relations between frame elements; and it is supplemented with sets of nouns that are compatible with the lexical units evoking the frame. A given conceptual frame is evoked by a set of lexical units, that (as of 2024) are exclusively defined for verbs.

Conceptual frames include a frame name, definition, semantic type, frame elements, and relations between frames. **Frame elements** possess a name, definition, semantic type, core status, and relations to each other, including **Core sets**, **Requires**, and **Excludes**. This information is inherited from the semantic frames via superframes, provided they are already defined, and then validated for Bulgarian through annotation.

Our rationale for employing conceptual frames and superframes is founded on the following arguments:

- **Not all lexical units evoking a given semantic frame manifest the same semantic structure, leading to varied syntactic behaviours.**

In FrameNet, lexical units are clustered based on sharing the same frame semantics, disregarding similarities in syntactic behaviour, unlike Levin's verb classes (Levin 1993), and within a FrameNet frame, there could be sets of verbs with related senses but distinct syntactic properties. Within the context of FrameNet's comprehensive approach to conceptual description, we strive to distinguish sets of lexical units with equivalent semantic properties. Therefore, we adhere to the principle that the semantic description of lexical units associated with a particular conceptual frame is achieved through the utilisation of the same number and type of frame elements. This approach maintains the structure of semantic frames, as multiple conceptual frames may correspond to one superframe, and through it to one semantic frame. Furthermore, a direct one-to-one correspondence between a FrameNet semantic frame and a conceptual frame is often absent due to differences in conceptualization across languages. Through the abstract superframe, conceptual frames describing a scene (either wholly or partially) are interconnected, both with each other and with the corresponding FrameNet semantic frame. For example, the Bulgarian verbs *вдявам* 'I am threading' and *вдяна* 'thread', with the definition in WordNet 'pass, pierce a thread or floor through the eye of a needle' evoke the **Placing** frame. This frame encompasses core elements such as **Agent**, **Theme**, and **Goal**. The **Agent** is part of a **Core set** with the frame element **Cause**, and either of them controls the **Theme**, positioning it at a location, the **Goal**. When describing the Bulgarian verbs *вдявам* and *вдя-*

на, only the **Agent** frame element holds relevance within the conceptual frame **Placing**, while the **Cause** frame element is omitted.

1.a. [Жената]_{Agent} ВДЯНА [конца]_{Theme} [в ухото на иглата]_{Goal}.
The woman inserted the thread into the eye of the needle.

1.b. *[Вятърът]_{Cause} ВДЯНА [конца]_{Theme} [в ухото на иглата]_{Goal}.
The wind inserted the thread into the eye of the needle.

• **Unlike English and other languages, a large part of diatheses in Bulgarian involve a lexical and/or morphological alteration of the base verb.**

In FrameNet, there is no specific list of verbal diatheses that a semantic frame encompasses. However, certain instructions in the annotation process imply that diatheses linked with a particular predicate are considered part of the frame associated with the basic diathesis. For instance, an additional frame is not created to accommodate uses like *Those boots sell well*, which depersonalises and generalises one or more of the prominent participants, such as the **Seller** (Ruppenhofer et al. 2016: 12). A similar rationale applies to passives. On the other hand, the systematic non-inheritance relationships between stative frames and the inchoative and causative frames referring to them are delineated using the frame-to-frame relations **Causative of** and **Inchoative of** (Ruppenhofer et al. 2016: 85).

In Bulgarian, several verbal diatheses exist, including se passives, impersonal participle passives, impersonal se passives, middles, anticausatives, and lexical reciprocals (Koeva 2022: 153). These diatheses can be classified into two categories: structure-preserving and structure-modifying. In structure-preserving diatheses, the number of frame elements remains constant, but at least one of the frame elements is changed (i.e., lexical reciprocals).

2.a. [Момчето]_{Author} ПИШЕ [на едно китайче]_{Addressee}.
The boy is writing to a Chinese.

2.b. [Момчето]_{Author₁}/Addressee₁ СИ ПИШЕ [с едно китайче]_{Author₂}/
Addressee₂.
The boy is corresponding with a Chinese.

Conversely, in structure-modifying diatheses, the number of frame elements differs across the alternating diatheses (impersonal passives), and these may be accompanied by a frame element's alternation (middles and anticausatives).

3.a. [Съседката]_{Grinder} ТРОШИ [лед]_{Patient} в кухнята.
The neighbour is crushing ice in the kitchen.

3.b. *Забранено е да СЕ ТРОШИ.*
Crashing is prohibited.

3.c. [Ледът]_{Theme} СЕ ТРОШИ лесно.
Ice crushes easily. 'The ice has the tendency to easily break.'

3.d. [Ледът]_{Theme} СЕ ТРОШИ.
Ice crushes. 'The ice has the tendency to break.'

Hence, the verbs *nuua* ('to compose a text in writing addressed to somebody') and *nuua cu* ('to correspond with somebody') will each be represented in distinct conceptual frames within the superframe **Text creation**. Similarly, the verbs *mpoua* ('to break into small pieces'), *mpouu ce* ('it breaks'), *mpouu ce* ('to have the tendency to easily break'), and *mpouu ce* ('to have the tendency to break') will be allocated to four separate conceptual frames within the superframe **Grinding**. This approach preserves the structure of the semantic frames without modification, while effectively reflecting the differences in both the semantic and syntactic structures of the verbs they represent.

Various types of information are encoded within the Bulgarian FrameNet. This includes lexical-semantic details such as lemma, part of speech, lexical type, semantic class, stylistic labels, and semantic relations. Additionally, grammatical information like verb aspect, transitivity, and the range of grammatical subjects is provided. Frame information encompasses frame definitions, frame-to-frame relations, frame elements, their definitions, core status, semantic types and relations, and ontological semantic classes of nouns suitable for pairing with a given target lexical unit. Furthermore, syntactic details for valency patterns are offered through the annotation of examples, including grammatical categories, grammatical functions, and the implicitness of frame elements. The sources for the information presented in Bulgarian FrameNet are outlined in Table 1.

| Type of information | Components | FrameNet | BulFrame |
|-------------------------|--------------------|----------------|-------------------|
| Lexical information | Lemma | FrameNet | WordNet, BulFrame |
| | Part of speech | FrameNet | WordNet |
| | Definition | FrameNet, OUP* | WordNet, BulFrame |
| | Semantic class | No | WordNet |
| | Stylistic note | No | WordNet |
| | Lexical type | FrameNet | BulFrame |
| | Semantic relations | No | WordNet |
| Grammatical information | Verb aspect | No | WordNet |
| | Transitivity | No | BulFrame |
| | Personality | No | BulFrame |

* OUP – Oxford University Press

| | | | |
|-----------------------|------------------------------|----------|--------------------|
| Frame information | Frame definition | FrameNet | FrameNet, BulFrame |
| | Frame-to-frame relations | FrameNet | FrameNet |
| | Frame elements | FrameNet | FrameNet, BulFrame |
| | Frame elements core status | FrameNet | FrameNet, BulFrame |
| | Frame elements definition | FrameNet | FrameNet, BulFrame |
| | Frame elements semantic type | FrameNet | FrameNet, BulFrame |
| | Frame elements relations | FrameNet | FrameNet, BulFrame |
| | Verb-to-Noun compatibility | No | BulFrame |
| Syntactic information | Grammatical category | FrameNet | BulFrame |
| | Grammatical function | FrameNet | BulFrame |
| | Implicitness | FrameNet | BulFrame |

Table 1. Sources of information in Bulgarian FrameNet

5.2 Noun frame elements fillers

Conceptual frames differ from semantic frames by linking frame elements to a set of lexical units for potential realisation. Each core frame element within a conceptual frame is associated with a set of nouns that are compatible with the verbs that evoke the frame (Koeva 2020: 17; Koeva 2021: 184 — 185). This set can consist of one, several or numerous nouns linked by semantic relations at the lexical level such as synonymy and antonymy or by hierarchical conceptual relations such as hypernymy and hyponymy. For example, the verb *варя* ‘boil’ with the definition ‘cook food in very hot or boiling water’ evokes the semantic frame **Apply heat**, which is described by four core frame elements: **Cook**, **Food**, **Container**, and **Heating instrument**. Each frame element is linked to synsets (one or more) from the Bulgarian WordNet, which unite (as roots of hypernym subtrees) the sets of appropriate nouns for collocations with the target verb.

Cook: eng-30-00007846-n: {person}

Food: eng-30-07555863-n: {food}; eng-30-07649854-n: {meat};
eng-30-07775375-n: {fish}; eng-30-07707451-n: {vegetable}

Container: eng-30-03990474-n: {pot}

Heating instrument: eng-30-03343560-n: {fire}; eng-30-03543254-n: {stove};
eng-30-08581699-n: {hearth}

6. Towards an Extended Ontology of Semantic Classes of Nouns

The development of the Extended Ontology of Semantic Classes of Nouns is motivated by the development of the conceptual frames: abstract semantic descriptions, part of the Bulgarian FrameNet, in which frame elements are associated with semantic classes of nouns encompassing sets of nouns appropriate for collocations with target verbal lexical units evoking conceptual frames. The Extended Ontology of Semantic Classes of Nouns is based on WordNet, implying that each concept in the ontology is linked to a synonym set from WordNet whenever possible (although not mandatory). It is evident that some classes serve to ensure compatibility among multiple lexical units. However, there are also classes introduced solely to describe restrictions on several, or even just one, lexical unit.

The accepted approach involves selecting either the highest-ranked concept or a combination of concepts (part of the Ontology), which means the highest-ranked synset or a combination of synsets from the (Bulgarian) WordNet that encompasses all suitable noun synsets for the fillers of the frame elements (Koeva 2020: 17; Koeva 2021: 184–185).

This approach streamlines the development of a robust training dataset for automatically labelling nouns and their semantic classes as frame element instances, thereby facilitating the annotation of valency patterns and the assignment of noun fillers to frame elements. In contrast, the manual annotation process in FrameNet, which entails identifying valency patterns and potentially extracting sets of noun fillers from annotated examples, is more labor-intensive. Furthermore, it lacks automatic classification of noun classes due to the limited size of sense-annotated corpora (with disambiguated senses), such as the Bulgarian sense-annotated corpus (Koeva 2012).

As previously noted, WordNet categorises nouns into broad semantic classes, which may not sufficiently reflect the semantic preferences of a diverse range of verbs. Moreover, multiple hypernymy in WordNet result from consolidating diverse taxonomic relations into a singular hypernymy. To address these challenges, we propose the following: a) linking WordNet synsets with more detailed ontological representations of noun semantic classes to enhance the selection of noun fillers; b) resolving multiple hypernymy within the WordNet structure; c) introducing additional semantic classes to specific synsets within hypernymy subtrees to facilitate precise selection.

6.1. Other Representations of Noun Fillers

Each of the mentioned resources (WordNet, CPA, and FrameNet) follows its own unique methodology. However, none of them aims to explicitly delineate the permissible and potential combinations of verbs and nouns realised in context, despite the assignment of semantic types regarding the noun fillers in PDEV and FrameNet.

Another manually crafted resource, VerbAtlas, offers comprehensive coverage of English verbs, defining prototypical argument structures for each cluster of WordNet synsets that build a semantically coherent frame (Di Fabio et al. 2019: 627). It also offers a limited set of explicit semantic roles, selectional preferences for the arguments in frames, and links to WordNet and BabelNet (Navigli et al. 2021). To address data sparsity concerns, VerbAtlas adopts VerbNet’s roles (Kipper et al. 2008), reducing them from 39 to 27 (in contrast to FrameNet’s practically unlimited number of frame elements). Selectional preferences in VerbAtlas are manually labeled from a set of 122 “macro-concepts” defined by WordNet synsets, whose hyponyms are expected to be probable candidates for the corresponding argument slot (Di Fabio et al., 2019: 630), employing a strategy akin to a previous algorithm-based approach (Agirre, Martinez, 2001).

The comparison in the interpretation of the verb *phone* in FrameNet and VerbAtlas shows some differences. In FrameNet, it evokes the semantic frame **Contacting** which is defined as “A **Communicator** (whose **Location** may be indicated) directs a **Communication** to an **Addressee** at a particular **Address**” together with the verbs: *cable.v*, *call in.v*, *call up.v*, *call.v*, *contact.v*, *e-mail.v*, *fax.v*, *get in touch.v*, *get through.v*, *mail.v*, *page.v*, *phone in.v*, *phone.v*, *radio.v*, *reach.v*, *ring up.v*, *ring.v*, *telegraph.v*, *telephone.v*, *telex.v*, *write in.v*, *write.v* (some nouns are also present). On the other hand, VerbAtlas frames it under **Communicate • Contact** which is defined as “Communicate with a place or person; establish communication with, as if by telephone” and embraces the following verbs: *phone²*, *telephone*, *call up*, *contact*, *get through*, *get hold of*; *drop a line*, *write*; *get in touch*, *connect*, *touch base*, *correspond*, *get*, *commune*; *call*, *call in*, *grunt-hoot*, *telepathise*, *telepathize*; *telecommunicate*, *pant-hoot*, *ping*, *rich out*; *network*, *cell phone*, *reticulate*, *e-mail*, *email*, *netmail*, *call back*, *raise* (explicit synonymy between verbs in VerbAtlas is borrowed from WordNet; other semantic relations are retrievable through WordNet).

In VerbAtlas, the semantic roles describing the frame **Communicate • Contact** are **Agent**, **Patient**, **Topic**, **Recipient**, and **Instrument**, as opposed to the five core frame elements within the semantic frame **Contacting** in FrameNet: **Communicator** (“The person that receives the message from the **Communicator**”), **Communication** (“The information that the **Communicator** wishes to impart to the **Addressee**, often to get a particular response”), **Topic** (“This is the subject matter to which the message pertains. It is typically expressed as a PP Complement headed by *about*”), **Address** “This frame element is used for a (metaphorical) place in a system of communication where **Communicators** and **Addressees** can access the system (*call at 555885*)”, and **Addressee** (“The person that receives the message from the **Communicator**”) and ten non-core. Some of the verbs in VerbAtlas such as *phone*, *telephone*, *call up*; *contact*, *get through*, *get hold of* are only defined with the semantic roles **Agent**, **Patient**, **Topic**, and **Recipient**, while some other verbs,

² The lexical units that appear in both resources are underlined.

such as *telepathise*, *telepathize*; *telecommunicate*; *cell phone*; *email*, are additionally defined with the semantic role **Instrument**. In FrameNet, the frame elements {**Communication**, **Purpose** (a peripheral frame element), **Topic**} and {**Address**, **Addressee**} form **Core sets**, which means that at least one member of the group or all members can appear.

It seems that the frame elements **Communicator** and **Addressee** correspond to the semantic roles **Agent** and **Patient**, since **Patient** is characterised as [individual]. The lack of definitions for semantic roles makes it difficult to determine which of the frame elements, **Communication** and **Topic** (or both), corresponds to the semantic role **Topic** despite the coincidence of the names. The semantic role **Instrument**, described as **Implicit** and **Shadow**, possibly corresponds to the peripheral frame element **Medium** (“The physical or abstract setting in which the message is conveyed”). And the semantic role **Recipient** remains unclear in relation to the role **Patient**.

A comparison between the selectional preferences and the semantic types shows that in VerbAtlas, the **Agent** and **Patient** are specified as [individual], [social group] and [facility] respectively, while in FrameNet, the corresponding frame elements **Communicator** and **Addressee** are classified as [Sentient]. Similarly, VerbAtlas specifies **Topic** and **Recipient** as [entity], while FrameNet assigns the semantic type [Communication] to the frame element **Communication** and does not specify the frame element [**Topic**]. The implicit semantic role **Instrument** is characterised by a single concept rather than selectional preferences defining a set. For example, for the lexical unit *call back*, the implicit **Instrument** is a *phone*, while for the lexical unit *cell phone*, it is a *mobile phone*, and so forth.

For both resources, a prediction of verb-noun combinations presents a challenge, regardless of whether manual or corpus-observing methods are used. This is due to the abstract level of semantic types and selectional preferences and the inherent difficulties in reconciling figurative but acceptable usage. Although the number of verbs described in VerbAtlas still exceeds the number of verbs in FrameNet, the semantic and syntactic information (by means of frame elements, definitions of semantic frames and frame elements, relations between frames and frame elements, valence patterns, and annotation of examples) in FrameNet is much more extensive and comprehensive.

In the Brazilian FrameNet, each core frame element is analysed based on the aspect of the scene it represents, resulting in the assignment of one or more frames to the frame element (Torrent et al. 2022: 4 – 5). Only frames representing events, states, attributes, and relations are eligible for frame element-to-frame relations. The information provided by the core frame element definition or semantic type is used to determine the type of concept it refers to (e.g., *people*, *location*, *event*) and the top-level frame that represents it. By linking the conceptual structures that build it, such as semantic frames and frame elements, the FrameNet gains extra semantic information. The technique looks to be similar to the establishment of morphosemantic relations between verb and noun synsets in WordNet; however, this extension is not applicable to Bulgarian FrameNet because nouns have yet to be introduced in it.

Based on the review of the presented resources, several conclusions can be drawn. The definition of noun fillers for frame elements relies on ontological representations of abstract entities, some sourced from WordNet or linked to it. However, due to the relatively high level of abstraction (with only 253 semantic types used to classify nouns in CPA), accurately predicting noun fillers for frame elements (including semantic types for verb patterns' slots and selectional preferences for semantic roles) faces significant challenges.

The generalised concepts in the ontological representation of semantic classes of nouns to encompass a wide range of words, some of which are semantically incompatible with the target lexical units (apart from the fact that it is not technically possible to retrieve the members of the semantic classes unless they are linked to an extended ontology or to WordNet). Moreover, the WordNet noun hierarchies are constructed for a different purpose, and a hypernymy-hyponymy subtree may contain synsets belonging to different semantic classes or to different ontological classes, i.e., concrete and abstract nouns.

6.2. Mapping existing ontological representations to WordNet noun hierarchy

The extension of WordNet's 25 semantic classes includes the linking of WordNet concepts to different hierarchies (ontologies): CPA semantic types, FrameNet semantic types, and VerbAtlas selectional preferences.

Such a sub-classification has already been achieved by a manual mapping in which the semantic types of CPA were matched with the corresponding WordNet synsets (Koeva et al. 2018a). As a result, the synsets are categorised into 253 semantic CPA types in addition to the WordNet semantic classes, with hyponym noun synsets inheriting both the semantic class and the semantic type of their parent. The taxonomic organisation of WordNet facilitates inheritance between semantic classes and semantic types along the hierarchy, ensuring a more precise delineation of verb-noun compatibility.

For example, the synset `eng-30-07881800-n: {beverage; drink}` is marked both with the WordNet semantic class `noun.food` and CPA semantic type `[Beverage]`, which is inherited by its hyponyms `{smoothie}`, `{cider; cyder}`, `{wine}`, etc.; the synset `{wine}` is additionally marked with the semantic type `[Wine]` complementary to semantic class `noun.food` and semantic type `[Beverage]`, which is inherited by its hyponyms `{dessert wine}`, `{Burgundy; Burgundy wine}`, etc. As the organisation is taxonomic, the semantic type `[Wine]` shows that the entity is also `[Beverage]` and `noun.food`; however, the specification is more narrow and excludes solid food and nonalcoholic beverages as well as other types of alcoholic drinks.

Initially, 199 semantic types were mapped to one WordNet concept, 39 semantic types mapped to two WordNet concepts, 12 semantic types mapped to three concepts, 2 semantic types mapped to four concepts, and 1 semantic type mapped to five concepts (Koeva et al. 2018a: 75 – 76). For instance, the semantic

type [Physical object] was mapped to synsets eng-30-00001930-n: {physical entity}, eng-30-00002684-n: {object; physical object}, and eng-30-00003553-n: {whole; unit}, while its child concept, the semantic type [Animate], was mapped to two synsets: eng-30-00004258-n: {living thing; animate thing} and eng-30-00004475-n: {organism; being}. Similarly, the next-level semantic type [Animal] was also mapped to two synsets: eng-30-00015388-n: {animal; animate being} and eng-30-01861778-n: {mammal; mammalian}. However, such one-to-many mappings violate ontological representation.

The ambiguity surrounding the initial decisions was resolved during the development of Bulgarian FrameNet by enriching conceptual frames with appropriate semantic classes that describe the realisations of frame elements. Further elaboration on the mapping revealed the exact mappings, for example: semantic type [Physical object] to synset eng-30-00002684-n: {object; physical object}, semantic type [Animate] to eng-30-00004258-n: {living thing; animate thing}, and semantic type [Animal] to eng-30-00015388-n: {animal; animate being} (Figure 2).

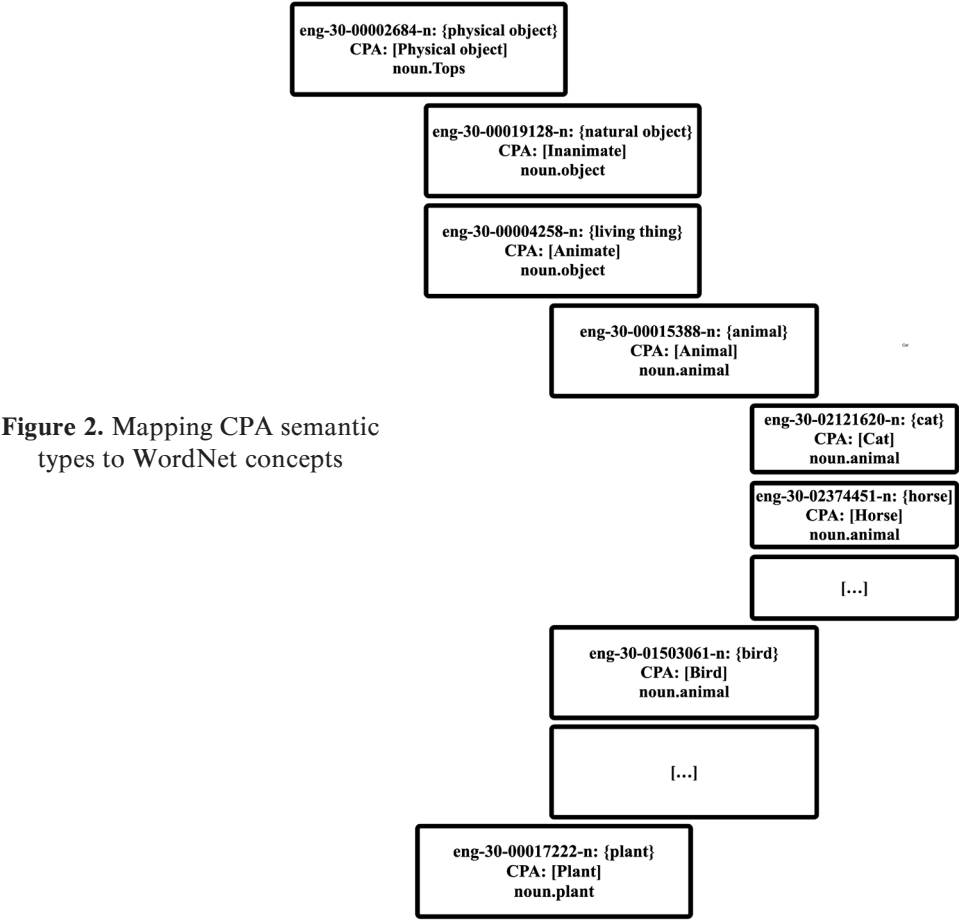


Figure 2. Mapping CPA semantic types to WordNet concepts

The small excerpt presented in Figure 2 also illustrates the sub-classification within the WordNet semantic classes, where the semantic class noun.animal was further categorised into semantic types such as [Cat], {Horse}, [Bird], and others.

Other ontological representations mapped to WordNet synsets include FrameNet semantic types and VerbAtlas selectional preferences. This forms the basic structure of the Extended Ontology of Semantic Classes of Nouns.

The hierarchy of semantic types in FrameNet (Ruppenhofer et al. 2016: 86) aligns with WordNet concepts and, consequently, with CPA semantic types (and VerbAtlas selectional preferences). Due to the abstract nature of semantic types in FrameNet and their alignment with WordNet, only four additional types are included in the Extended Ontology of Semantic Classes of Nouns, different from the CPA semantic types. For example, the semantic type [Line] is linked to the synset eng-30-08593262-n: {line} with the semantic class noun.shape. Throughout the mapping process, any terminological differences used to denote the same concepts were standardised.

The mapping process between the VerbAtlas selectional preferences and WordNet concepts was facilitated by VerbAtlas's explicit referencing to BabelNet synsets, which in turn are linked to WordNet synsets. A comparison between the semantic types in the CPA and the selectional preferences in VerbAtlas reveals extensive overlap, with only a few unique selectional preferences found in VerbAtlas but not in the CPA. For example, the selectional preference [Liquid] exists in both the CPA and the VerbAtlas, and has already been mapped to the WordNet synset eng-30-14940100-n: {liquid} with the semantic class noun.substance. Only seven selectional preferences are not present in the CPA, and they were additionally added in the Extended Ontology of Semantic Classes of Nouns. For example, the selectional preference [Shot] is mapped to the synset eng-30-00565302-n: {stroke, shot}.

The mapping of the three ontological representations to WordNet concepts is illustrated in Figure 3 with the same example. Despite the overlap between four concepts at a very abstract level, they provide a stable foundation for the Extended Ontology of Semantic Classes of Nouns. Three and two overlaps also offer a certain level of confidence.

However, some levels within the hierarchy of the Ontology remain unpopulated, as they will be filled in during the process of defining concrete lexical units, conceptual frames, and frame elements in the Bulgarian FrameNet.

To conclude, the initial development of the Extended Ontology of Semantic Classes of Nouns entails mapping the relevant ontologies to the WordNet noun hierarchy, allowing for the identification of nodes indicating suitable noun fillers for specific frame elements. Through this mapping process, the original 25 semantic classes in WordNet were initially enriched with an additional 253 semantic types from the CPA. The ambiguity in the initial mapping of CPA semantic types was resolved by relying on evidence from the Bulgarian FrameNet and further validated by mapping it to the selectional preferences of the VerbAtlas and the semantic types of the FrameNet. Further, the mapping of CPA semantic

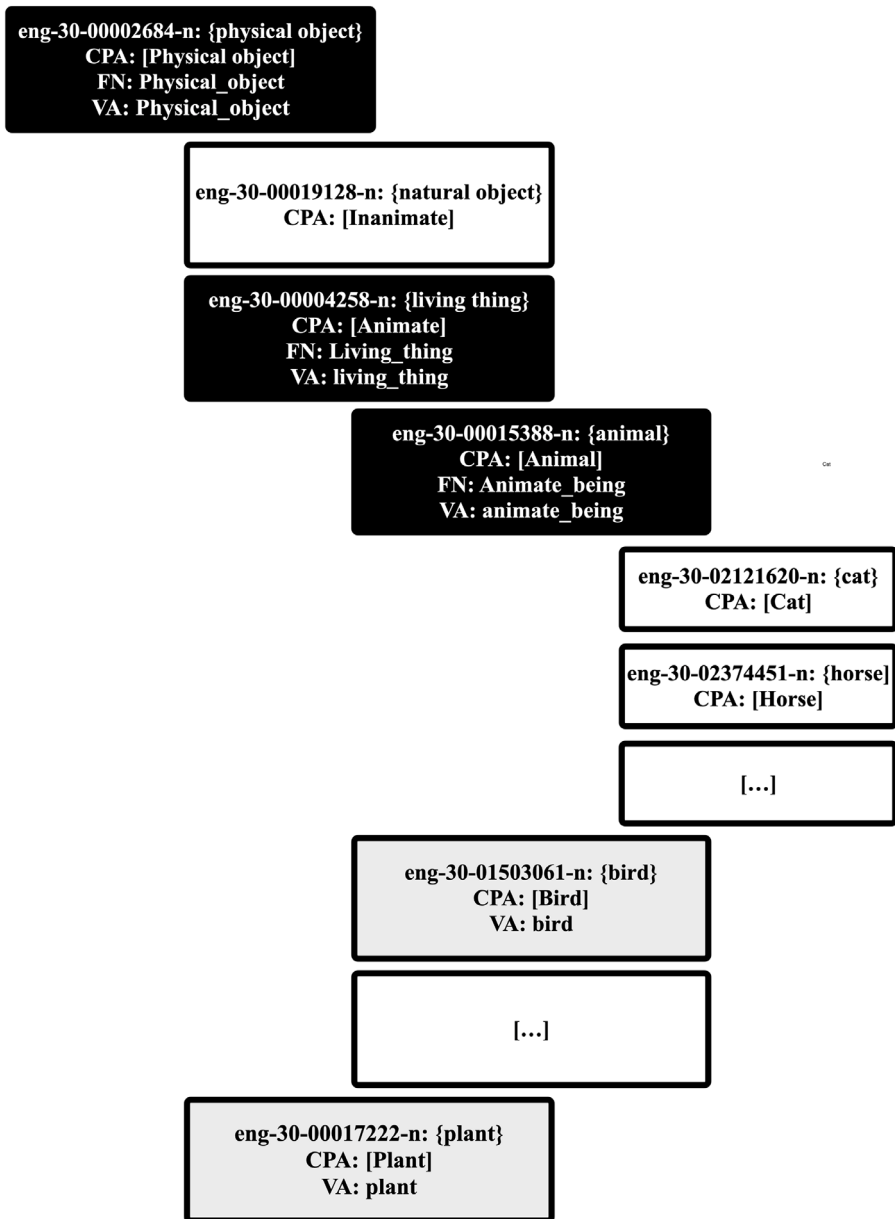


Figure 3. Mapping CPA semantic types, FrameNet semantic types, and VerbAtlas selectional preferences to WordNet concepts

types is augmented by incorporating an additional 7 selectional preferences from VerbAtlas and 4 semantic types from FrameNet.

The Extended Ontology of Semantic Classes of Nouns can be further expanded with new, more specific classes if the number of annotated Bulgarian verbs and elaborated frame elements increases, necessitating additional characterisation.

6.3. Resolving multiple hypernymy

In WordNet, the hypernymy relation may encompass various sub-relations, leading to hierarchies where nouns with vastly different characteristics coexist. For example, abstract and concrete nouns may appear together in a WordNet noun subtree, implying compatibility with verbs whose frame elements accommodate a broad range of nouns, such as *I think of* [entity: {*gesture*}; {*idea*}]. Some other verbs require nouns from specific classes, as seen in *I see* [physical object: {*gesture*}] *[abstraction: {*idea*}], and the inheritance of classes from multiple hypernyms introduces ambiguity. Therefore, to ensure the unambiguous inheritance of noun semantic classes, it is necessary to eliminate instances of multiple hypernymy and restrict class inheritance solely to the **is-a** relation or true hypernymy.

The structure of nouns in WordNet forms a directed connected graph comprising various semantic relations, with the taxonomic relation **is-a** (hypernymy) and its inverse relation (hyponymy) being most important for the semantic organisation. Both relations form a tree where noun synsets are interconnected through unique paths, with hypernyms having multiple hyponyms and each hyponym linked to exactly one hypernym.

Multiple hypernyms are classified into three types: exclusive, conjunctive, or non-exclusive, as outlined by EAGLES (1999). Exclusive multiple hypernyms (*albino* is either an *animal* or a *human*), often associated with polysemy, assign different hypernyms to distinct meanings of the same word. In contrast, conjunctive multiple hypernyms combine various semantic relations (*spoon* is both *cutlery* and *container*). Non-exclusive hypernyms accommodate both disjunctive and conjunctive relations (*knife* can be *cutlery*, a *weapon* or both), but the last ones are generally avoided in WordNet since different concepts should not be encoded within the same synset (node).

The current version of WordNet includes 1,421 synsets featuring multiple hypernyms, thereby undermining its taxonomic structure and the corresponding ontological representation of noun semantic classes. Figure 4 provides an illustration of three concepts characterised by multiple hypernyms. Moreover, the synsets depicted in Figure 4 (in grey) serve as examples of the alterations in noun semantic classes within certain noun hypernym sub-trees.

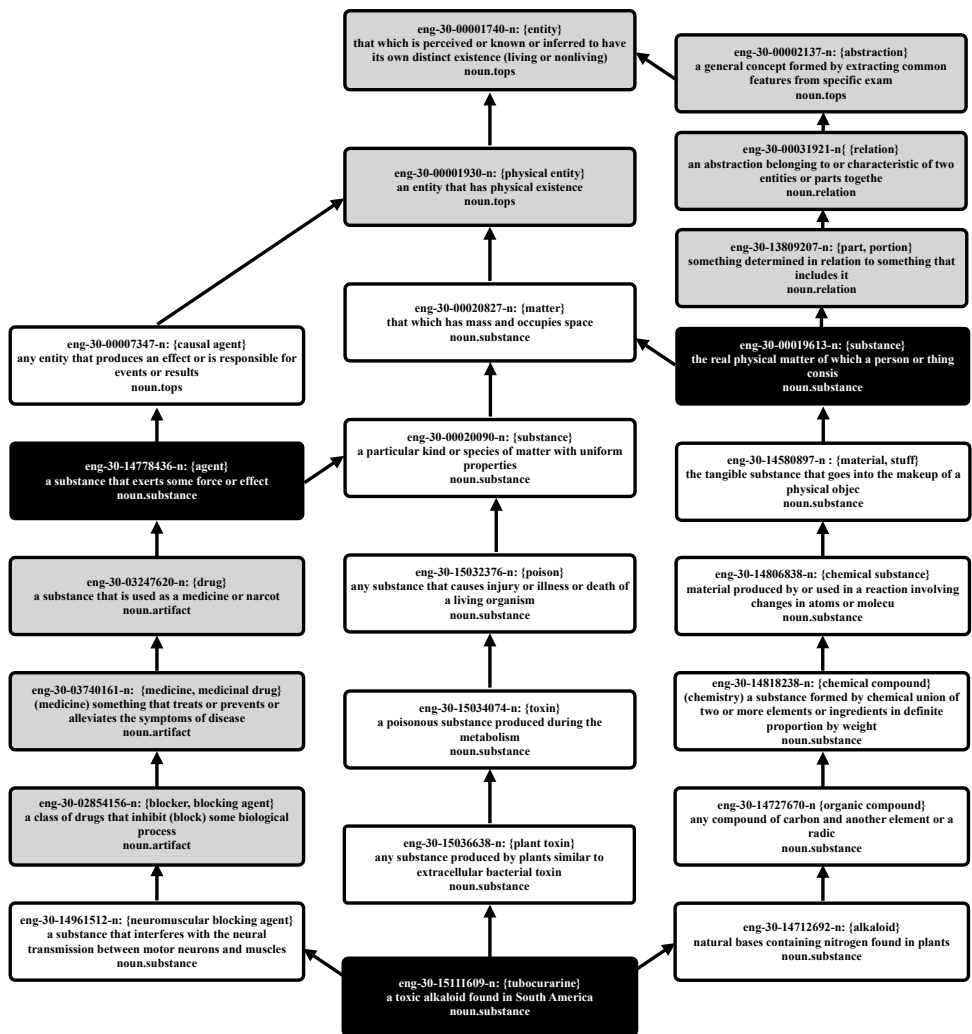


Figure 4. Multiple hypernyms in WordNet

Drawing on the principle that a synset ought to be linked to a solitary hypernym, the issue of multiple hypernymy was addressed through one of the following approaches (Koeva, Hristov 2023: 349) (Figure 5):

- Converting a multiple hypernymy relation to one of nine alternative relations: **origin**, **form**, **function**, **characteristic**, **purpose**, **use**, **member**, **part**, or **portion**, the first three of which have already been proposed (Koeva et al. 2018b) and the last four of which are used in WordNet;
- Removal of a hypernymy relation when it lacks appropriate connectivity (rarely);

Figure 5 presents the decisions on resolving multiple hypernymy.

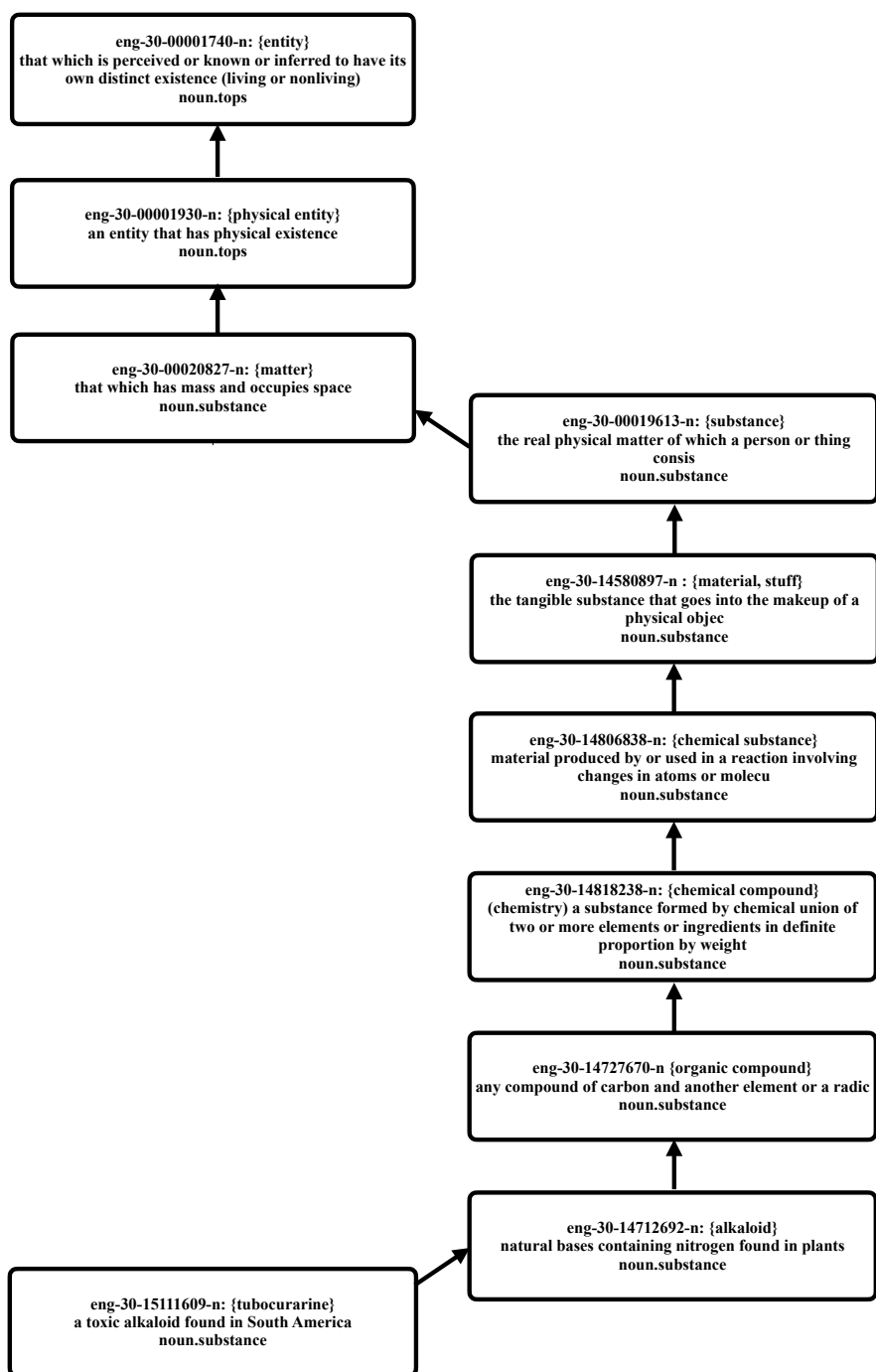


Figure 5. Resolving multiple hypernymy

- Introduction of a new hypernymy relation if none of the currently linked hypernyms are deemed suitable (rarely).

For example, the synset eng-30-15111609-n: {tubocurarine} ‘a toxic alkaloid found in South America’ was originally associated with three hypernyms: eng-30-14961512-n: {neuromuscular blocking agent} ‘a substance that interferes with the neural transmission between motor neurons and skeletal muscles’; eng-30-15036638-n {plant toxin, phytotoxin} ‘any substance produced by plants that is similar in its properties to extracellular bacterial toxin’, and eng-30-14712692-n: {alkaloid} ‘natural bases containing nitrogen found in plants’. Following the analysis, the relation with {neuromuscular blocking agent} was reclassified as **function**, while the relation with eng-30-15036638-n: {plant toxin, phytotoxin} was designated as **purpose**.

Similarly, the synset eng-30-00019613-n: {substance} ‘the real physical matter of which a person or thing consists’ initially is linked to two hypernyms: eng-30-13809207-n: {part; portion; component} ‘something determined in relation to something that includes it’ and eng-30-00020827-n: {matter} ‘that which has mass and occupies space’. In this case, the relation to {part; portion; component} was removed.

This approach allows the redefinition of certain hypernymy relations and resolves instances of multiple hypernyms. By delineating distinct semantic relations, we can utilise solely the **is-a** inheritance relation to categorise the semantic classes of noun synsets into more refined groups and anticipate the compatibility of verbs and nouns.

6.4. Identification of noun fillers based on corpus analysis

For the set of verbs evoking a particular conceptual frame, suitable examples are sought in the Bulgarian National Corpus³ or in other sources. By annotating examples that show the syntactic realisation of the frame elements, the valence patterns for a given lexical unit are constructed. The collection of examples also serves to observe which nouns are suitable for the syntactic realisation of the frame elements (as a single noun, noun phrase, or prepositional phrase).

This task is conducted by annotators, who select WordNet nodes that could potentially represent the set of candidate fillers and verify whether the concept is already included in the Extended Ontology of Semantic Classes of Nouns. Subsequently, they assess whether the subordinate noun synsets are suitable for pairing with the target verb as the syntactic realisation of its frame elements and make a decision on whether to utilise a single concept from the Ontology, a set of concepts, or introduce a new concept for defining the noun fillers.

To illustrate the proposed approach, the noun fillers for the frame elements of the **Statement** frame evoked by the verbs *обяснявам* and *обясня* (explain) ‘make plain and comprehensible’ will be presented. The core frame elements are

³ <https://search.dcl.bas.bg>

defined in FrameNet as follows: “The **Speaker** is the sentient entity that produces the **Message** (whether spoken or written); The **Message** is the frame element that identifies the content of what the **Speaker** is communicating to the **Addressee**. It can be expressed as a clause or as a noun phrase; The **Topic** is the subject matter to which the **Message** pertains. It is normally expressed as a PP Complement headed by *about*, but in some cases it can appear as a direct object”. For example, *Evelyn spoke candidly about her past*; “**Medium** is the physical entity or channel used by the **Speaker** to transmit the statement”. For example, *Reports say Iran is working on laser enrichment technologies*. There are several non-core frame elements, among which **Addressee** is defined as: “The **Addressee** is the person to whom the **Message** is communicated. When this frame element is expressed, it often appears in a prepositional phrase introduced by *to*, or as a direct object”.

The noun fillers for the frame element **Speaker** are either nouns that are assigned to the semantic class *noun.person* in WordNet or non-sentient nouns whose meaning can express unions of persons, such as *party, ministry, organisation, company*, etc., denoting organisations responsible for certain functions, policies, or services. In this context, such nouns embody abstract concepts of administrative authorities, policy formulation, regulatory oversight, etc., which refer not to physical, tangible entities but to the collective functions and responsibilities associated with human activities. The concepts in the Ontology are defined by: a) the WordNet synset *eng-30-02472293-n*: {human, human being}, which matches the semantic type [Human] in CPA and FrameNet (although the frame element is specified as [Sentient]) and the selectional preference [human] in VerbAtlas; b) the WordNet synset *eng-30-08008335-n* {organisation}, which matches the semantic type [Institution] in CPA and [Organisation] in FrameNet and the selectional preference [social group] in VerbAtlas. This selectional preference corresponds to the synset *eng-30-07950920-n*: {social group} ‘people sharing some social relation’, which, however, is overly abstract. Among its direct hyponyms are synsets such as: *eng-30-07966140-n*: {society} ‘an extended social group having a distinctive cultural and economic organisation’; *eng-30-07966719-n*: {sector} ‘a social group that forms part of the society or the economy’; etc. Therefore, the most appropriate direct hyponym {organisation} was chosen as part of the Ontology. The same criterion was applied to assign the Ontology semantic classes to the frame element **Addressee**.

The primary consideration regarding the frame element **Message** fillers is that they ought to be nouns categorised as either *noun.communication* or *noun.cognition* in WordNet. However, these nouns differ in how they express communication and cognition. Therefore, it is important to develop a technique to eliminate the nouns that cannot be collocated with the verb *explain* as objects.

The synset *eng-30-00033020-n*: {communication}, defined as ‘something that is communicated by or to, or between people or groups’, stands at the root of the hierarchy of nouns classified under the semantic class *noun.communication*. Nevertheless, this concept is too abstract to serve as a filler for the **Message** frame element.

Although the hyponyms of this synset (direct and full) are appropriate in some cases, others introduce inappropriate nouns to be combined with the target verbs. For example: eng-30-06520222-n: {receipt} ‘an acknowledgment (usually tangible) that payment has been made’; eng-30-06275634-n: {mail} ‘the bags of letters and packages that are transported by the postal service’; and eng-30-01102436-n: {publication} ‘the communication of something to the public; making information generally known’, among others.

Some of the non-combinable nouns within the subtree are concrete nouns {receipt} and {mail}, while others are labelled with the semantic class noun.act as {publication}. In such cases, an appropriate approach is to narrow down the set to nouns classified under the semantic class noun.communication and introduce another layer of classification: abstract and concrete nouns.

The synset at the top of the hierarchy, designated with the semantic class noun.cognition, is eng-30-00023271-n: {cognition; knowledge}, defined as ‘the psychological result of perception, learning, and reasoning’. However, not all of its hyponyms are suitable as fillers for the **Message** frame element. To address this issue, two more specific concepts are chosen: eng-30-05816287-n {information} ‘knowledge acquired through study or experience or instruction’ and eng-30-05833840-n {idea; thought} ‘the content of cognition; the main thing you are thinking about’. Both of these concepts are also defined within the selectional preferences in VerbAtlas, and the former is included in the CPA semantic types as well.

For the frame element **Topic**, the concept represented by the synset eng-30-00002137-n {abstraction; abstract entity}, defined as ‘a general concept formed by extracting common features from specific examples’, is chosen.

Medium can be conveyed through hyponyms of several concepts presented by the synsets: eng-30-06722453-n: {writing; written material}; eng-30-06722453-n: {statement}; eng-30-06263369-n: {press; public press}; eng-30-06277280-n: {television; telecasting}; and eng-30-06619428-n: {broadcast; program}, all falling under the semantic class noun.communication. The corresponding semantic types in CPA are [Television program] and [Document].

It can be concluded that the accurate determination of the appropriate noun classes to fill the positions of the frame elements of a given verb necessitates the combination of two approaches:

Selecting the most suitable concept or combination of concepts from the Extended Ontology of Semantic Classes of Nouns (presented by the WordNet noun synsets that dominate appropriate nouns).

Introducing additional elementary semantic types and classifying WordNet noun synsets based on these types to provide correct generalisations. These types may encompass *collective*, *abstract*, *concrete*, *agentive* and so on.

The development’s contribution lies in detailing the systematically conceived and executed steps to establish a suitable set of semantic classes organised in the Extended Ontology of Semantic Classes of Nouns. The creation of the Ontology of Semantic Classes of Nouns was guided by the following principles:

a) Utilising available ontological representations of semantic noun classes, comparing and incorporating both overlapping and unique concepts based on real examples used to develop conceptual frames for Bulgarian FrameNet.

b) Aligning ontology concepts with synonym sets from WordNet to leverage WordNet's taxonomic organisation and associate ontology concepts with sets of nouns that are their hyponyms in WordNet.

c) Validation and refining the taxonomic organisation of nouns in WordNet to resolve instances of multiple hypernymy.

d) Augmenting the taxonomic organisation of nouns in WordNet with additional noun classifications, including divisions into abstract and concrete, animate and inanimate, human and non-human, agentive and non-agentive, where distinctions are not evident from existing semantic classes.

The first principle draws upon the ontology of semantic types in CPA and FrameNet, as well as the ontological representation of selectional preferences in VerbAtlas. While theoretically, these ontological representations can be supplemented with new concepts, they are in practice fixed structures. For example, the CPA ontology allows for the inclusion of new concepts based on new verb patterns; however, the resource has remained stagnant for many years. Moreover, the number of verb lexical units in FrameNet (over 5000), VerbAtlas (over 13 000), and CPA (approximately 1700 words with more than 5000 verb patterns) underscores the need for a comprehensive approach to model the compatibility of varied meanings of lexical units in context.

The second principle is followed to some extent in VerbAtlas and to a lesser extent in FrameNet, but both resources provide only a high-level representation of semantic combinations between verbs and nouns, akin to dictionaries.

Regarding the third and fourth principles, there have been no known attempts to address multiple hypernymy in WordNet or supplement its concepts with new semantic classes to distinguish between concrete and abstract nouns, animate and inanimate, etc.

7. Conclusions

The advancement of modern technologies and the emergence of powerful large language models have greatly enhanced the ability to predict the next word in a given context, particularly in English but increasingly in other languages as well. This development prompts the question of whether there is a need for classifying noun classes based on their compatibility with verbs. According to the authors of this study, the necessity for such classification is justified by the following factors:

Firstly, the proposed approach for classification is primarily useful for in-depth analysis and study of a particular language, whether it be one's native tongue or a foreign language. The classification of nouns not only illustrates word compatibility but also organises nouns based on shared semantic properties, offering insights into their ontological and independent class groupings.

Secondly, such classification can enhance the fine-tuning and performance of large language models by enabling them to accurately predict the compatibility of verbs and nouns not only in common contexts but also in rare and unconventional ones. This refinement can lead to more nuanced and contextually appropriate language generation.

The ultimate goal is to identify classification models that effectively describe noun classes, whether achieved by experts, automatically, or through a combination of both methods. This approach aims to enhance our understanding of language structure and improve the capabilities of language processing systems.

Acknowledgments: This study is carried out as part of the scientific programme under the project *Enriching the Semantic Network WordNet with Conceptual Frames* funded by the Bulgarian National Science Fund (Grant Agreement No. KP-06-N50/1 of 2020).

References

- Agirre, Martinez 2001: Agirre, E., D. Martinez. Learning class-to-class selectional preferences. – In: *Proceedings of the ACL 2001 Workshop on Computational Natural Language Learning (ConLL)*. <<https://aclanthology.org/W01-0703>> [20.01.2024]
- Basile 2015: Basile, V. WordNet as an Ontology for Generation. – In: *WebNLG 2015 1st International Workshop on Natural Language Generation from the Semantic Web, Jun 2015, Nancy, France*. <<https://inria.hal.science/hal-01195793/document>> [20.01.2024]
- Borgo et al. 2022: Borgo, S., R. Ferrario, A. Gangemi, N. Guarino, C. Masolo, D. Porello, E. M. Sanfilippo, L. Vieu. DOLCE: A Descriptive Ontology for Linguistic and Cognitive Engineering. – *Applied Ontology*, 17 (1), pp. 45 – 69.
- Cinkova, Hanks 2010: Cinkova, S., P. Hanks. Validation of Corpus Pattern Analysis. – In: *Assigning Pattern Numbers to Random Verb Samples*, pp. 1 – 16. <https://ufal.mff.cuni.cz/project/spr/data/publications/annotation_manual.pdf> [20.01.2024]
- Di Fabio et al. 2019: Di Fabio, A., S. Conia, R. Navigli. VerbAtlas: a novel large-scale verbal semantic resource and its application to semantic role labeling. v In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong: Association for Computational Linguistics, pp. 627 – 637. <<https://aclanthology.org/D19-1058/>> [20.01.2024]
- EAGLES 1999: *Preliminary Recommendations on Lexical Semantic Encoding. Final report*. <<https://www.ilc.cnr.it/EAGLES96/EAGLESLE.PDF>> [20.01.2024]
- Fellbaum 1990/1993: Fellbaum, C. English Verbs as a Semantic Net. – *International Journal of Lexicography*, Volume 3, Issue 4, Winter 1990, pp. 278 – 301; reprinted in 1993, pp. 40 – 51. <<https://wordnetcode.princeton.edu/5papers.pdf>> [20.01.2024]
- Fillmore 1976: Fillmore, C. J. Frame semantics and the nature of language. – *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, 280(1), pp. 20 – 32.
- Fillmore 1982: Fillmore, C. J. Frame semantics. – In: *Linguistics in the morning calm*. Seoul: Hanshin Publishing, pp. 111 – 137.
- Fillmore 2007: Fillmore, C. J. Valency issues in FrameNet. – In: Herbst, T., K. Gutz-Votteler (Eds.). *Valency: Theoretical, descriptive and cognitive issues*. Berlin: Mouton de Gruyter, pp. 129 – 160.

- Fillmore, Baker 2010: Fillmore, C. J., C. Baker. A frames approach to semantic analysis. – In: Heine, B., H. Narrog (Eds.). *The Oxford handbook of linguistic analysis*. Oxford: Oxford University Press, pp. 313 – 340.
- Fillmore et al. 2003: Fillmore, C. J., C. R. Johnson, M. R. L. Petrucci. Background to FrameNet. – *International Journal of Lexicography*, 16(3), pp. 235 – 250.
- Gangemi et al. 2003: Gangemi, A., R. Navigli, P. Velardi. The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet. – In: Meersman, R., Z. Tari, D. C. Schmidt (Eds.). *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE. OTM 2003. Lecture Notes in Computer Science*, vol 2888. Berlin, Heidelberg: Springer, pp. 820 – 838.
- Hanks 2004: Hanks, P. Corpus pattern analysis. – In: *Proceedings of the 11th EURAL- EX International Congress (EURALEX 2004), Lorient, France, 6–10 July 2004*. Lorient: Université Bretagne Sud, pp. 87 – 97.
- Hanks 2012: Hanks, P. How people use words to make meanings: Semantic types meet valencies. – In: Boulton, A., J. Thomas (Eds.). *Input, Process and Product: Developments in Teaching and Language Corpora*. Brno: Masaryk University Press, pp. 54 – 70. <<https://www.cceol.com/search/chapter-detail?id=837374>> [29.02.2024]
- Hanks, Pustejovsky 2005: Hanks, P., J. Pustejovsky. A Pattern Dictionary for Natural Language Processing. – *Revue française de linguistique appliquée*, 2 (Vol. X), pp. 63 – 82. <<https://www.cairn.info/revue-francaise-de-linguistique-appliquee-2005-2-page-63.htm>> [20.01.2024]
- Koeva 2010: Koeva, S. Lexicon and grammar in Bulgarian FrameNet. – In: *Proceedings of the seventh international conference on language resources and evaluation (LREC'10)*. Valletta: European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2010/pdf/705_Paper.pdf> [20.01.2024]
- Koeva 2012: Koeva, S. Bulgarian Sense-annotated Corpus – between the tradition and novelty. – In: *Études Cognitives | Studia Kognitywne*, 12, pp. 181 – 198.
- Koeva 2020: Koeva, S. Semantic Relations and Conceptual Frames. – In: Koeva, S. (Ed.). *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Institute for Bulgarian Language Prof. Lyubomir Andreychin, pp. 7 – 20.
- Koeva 2021: Koeva, S. Towards Expanding WordNet with Conceptual Frames. – In: *Proceedings of the 11th Global WordNet Conference*. University of South Africa (UNISA): Global WordNet Association, pp. 182 – 191.
- Koeva 2022: Koeva, S. The system of diatheses in Bulgarian and stative predicates. – In: Koeva, S., E. Ivanova, Y. Tisheva, A. Zimmerling (Eds.). *Ontology of Stative Situations – Linguistic Modeling. A Contrastive Bulgarian-Russian Study*. Sofia: Prof. Marin Drinov Publishing House of the Bulgarian Academy of Sciences, pp. 117 – 160.
- Koeva, Hristov 2023: Koeva, S., D. Hristov. Resolving Multiple Hyperonymy. – In: *Proceedings of the 12th Global WordNet Conference*. University of the Basque Country, Donostia – San Sebastian, Basque Country: Global WordNet Association, pp. 343 – 351.
- Koeva et al. 2003: Koeva, S., E. Doychev, G. Cholakov. Formal representation of syntactic environment and syntactic features of words (SYNText – a WEB base system designed for frame lexicon). – In: *Balkan language resources and tools*, pp. 41 – 47.
- Koeva et al. 2008: Коева, С., П. Влахова, Р. Декова, П. Несторова, А. Атанасов. Представяне на езиковата информация в Българския ФреймНет – лингвистична мотивировка. – В: *Българският ФреймНет. Семантико-синтактичен речник на българския език*.

- Коева, С. (съст.). София: Институт за български език „Проф. Любомир Андрейчин“, с. 58 – 78. (Коева, S., R. Vlahova, R. Dekova, P. Nestorova, A. Atanasov. Predstavuane na ezikovata informatsia v Balgarskiya FreymNet – lingvistichna motivirovka. – In: *Balgarskiyat FreymNet. Semantiko-sintaktichen rechnik na balgarskiya ezik*. Koeva, S. (sast.). Sofia: Institut za balgarski ezik „Prof. Lyubomir Andreychin“, pp. 58 – 78.)
- Koeva et al. 2018a: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Mapping WordNet concepts with CPA ontology. – In: *Proceedings of the 9th Global WordNet Conference (GWC'2018)*. Singapore: Global WordNet Association, pp. 70 – 77.
- Koeva et al. 2018b: Коева, Св., В. Стефанова, Д. Христов. Семантични релации в рамките на многократната хиперонимия в УърдНет. – *Чуждоезиково обучение*, 45, 4, 2018, с. 343 – 365. (Коева, S., V. Stefanova, D. Hristov. Semantichni relacii v ramkite na mnogokratnata hiperonimiya. – *Chuzhdoezikovo obuchenie*, 45, 4, 2018, pp. 343 – 365.)
- Laparra et al. 2012: Laparra, E., G. Rigau, P. Vossen. Mapping WordNet to the Kyoto ontology. – In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. Istanbul: European Language Resources Association (ELRA), pp. 2584 – 2589.
- Loenneker-Rodman, Baker 2009: Lönneker-Rodman, B., C. F. Baker. The FrameNet model and its applications. – *Natural Language Engineering*, 15(3), pp. 415 – 453.
- Miller 1986: Miller, G. Dictionaries in the mind. – *Language and Cognitive Processes*, 1, pp. 171 – 185.
- Miller 1990/1993: Miller, G. Nouns in WordNet: A Lexical Inheritance System. – *International Journal of Lexicography*, 3(4), 1990, pp. 245 – 264; reprinted in 1993, pp. 10 – 25. <<https://wordnetcode.princeton.edu/5papers.pdf>> [20.01.2024]
- Miller et al. 1990/1993: Miller, G., R. Beckwith, C. Fellbaum, D. Gross, K. Miller. Introduction to WordNet: an On-line Lexical Database. – *International Journal of Lexicography*, 3(4), 1990, pp. 233 – 244; reprinted in 1993, pp. 1 – 9. <<https://wordnetcode.princeton.edu/5papers.pdf>> [20.01.2024]
- Navigli et al. 2021: Navigli, R., M. Bevilacqua, S. Conia, D. Montagnini, F. Cecconi. Ten Years of BabelNet: A Survey. – In: *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence Survey Track*. pp. 4559 – 4567. <<https://www.ijcai.org/proceedings/2021/0620.pdf>> [20.01.2024]
- Niles, Pease 2003: Niles, I., A. Pease. Linking lexicons and ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. – In: *Proceedings of the 2003 International Conference on Information and Knowledge Engineering*. Las Vegas, Nevada. <<https://www.adampease.org/Niles-IKE.pdf>> [20.01.2024]
- Pustejovsky et al. 2004: Pustejovsky, J., P. Hanks, A. Rumshisky. Automated Induction of Sense in Context. – In: *COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics*. Geneva: COLING, pp. 924 – 930. <<http://acl.ldc.upenn.edu/C/C04/C04-1133.pdf>> [20.01.2024].
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: extended theory and practice*. Berkeley, California: International Computer Science Institute.
- Sikos, Pado 2018: Sikos, J., S. Pady. Framenet's using relation as a source of concept-based paraphrases. – *Constructions and Frames*, 10(1), pp. 38 – 60.
- Torrent et al. 2022: Torrent, T. T., E. E. da Sylva Matos, F. Belcavello, M. A. Viridiano, M. Gamonal, A. D. da Costa, M. C. Marim. Representing Context in FrameNet: A Multidimensional,

РАЗШИРЕНА ОНТОЛОГИЯ НА СЕМАНТИЧНИТЕ КЛАСОВЕ НА СЪЩЕСТВИТЕЛНИТЕ ИМЕНА

Светла Коева

Институт за български език „Проф. Любомир Андрейчин“,
Българска академия на науките
svetla@dcl.bas.bg

Резюме. Основната цел на изследването е да се предложи ефективен подход за характеризирание на фреймовите елементи от концептуалните фреймове на Българския фреймнет посредством класове съществителни имена, с цел да се предложат изчерпателни и семантично валидни комбинации между глаголните лексикални единици, предизвикващи представата за концептуален фрейм, и подходящи съществителни. Студията предлага кратък преглед на семантичните класификации на съществителните в Уърднет, Корпусния анализ на изреченските модели, Фреймнет и Върбатлас, като се подчертава тяхното значение за определяне на съчетаемостта между глаголи и съществителни. Представя се структурата на Българския фреймнет (Булфрейм), който съдържа валидните за български компоненти, заимствани от Фреймнет, но и голямо количество лексикална, морфологична, синтактична и семантична информация, специфична за български език. Една от отличителните характеристики на Българския фреймнет е спецификацията на класове съществителни имена, които показват множеството от подходящи съществителни за лексикална реализация на фреймовите елементи. Чрез съотнасяне на синонимни множества от Уърднет със семантичните типове на Корпусния анализ на изреченските модели и Фреймнет, както и със селективните предпочитания на Върбатлас, се изгражда основната структура на Разширената онтология на семантичните класове на съществителните. Понятията в Онтологията са свързани (не изключително) със синонимните множества от Уърднет и следователно, с множества от съществителни, подходящи за свързване с глаголните лексикални единици, които предизвикват представата за съответните концептуални фреймове. Приносът на разработката се състои в детайлизиране на стъпките за подбор на семантичните класове, изграждащи Разширената онтология на семантичните класове на съществителните.

Ключови думи: *семантичен клас, семантичен тип, селективни предпочитания, семантичен фрейм, фреймов елемент*

Svetla Koeva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52 Shipchenski prohod Blvd., bl. 17
Sofia 1113, Bulgaria
<https://orcid.org/0000-0001-5947-8736>

ЧАСТ НА ТЯЛОТО КАТО СЕМАНТИЧНА ХАРАКТЕРИСТИКА

Валентина Стефанова

Институт за български език „Проф. Любомир Андрейчин“

valentina@dcl.bas.bg

Резюме. Студията се фокусира върху анализ на фреймовия елемент **Част на тялото (Body part)** като семантична характеристика за описание на семантични фреймове според йерархичната структура във Фреймнет. Целта е семантично описание и представяне на методологията за извличане на информация и анализ на група глаголи според наличие на конкретен фреймов елемент. Изследването е част от създаването на Българския фреймнет (Булфрейм) – система за визуализация и редакция на концептуални фреймове, които са дефинирани в системата (Коева, Douchev 2022). Анализът се осъществява чрез извличане на глаголи, които се описват концептуално чрез семантични фреймове с ядрен фреймов елемент **Част на тялото**. Използвани са семантични ресурси, създадени за целите на семантичното, синтактичното и концептуалното знание за глаголни единици. Направено е съотнасяне с глаголни синонимни множества от Българския уърднет. Съществителните имена, които глаголет присъединява, се разглеждат с оглед на отделни синонимни множества и семантични типове като възможна реализация на глагола. Като резултат е предложено групиране на съществителни имена чрез асоциирани семантични типове от Корпусния анализ на моделите (КАМ).

Ключови думи: *семантичен фрейм, фреймов елемент, семантичен тип, част на тялото*

1. Въведение

Правени са множество сравнителни изследвания, свързани с обозначаване на частите от човешкото тяло в различни езици (Andersen 1978; Brown 1976,; Enfield 2006; Enfield et al. 2006; Wierzbicka 2006 и др.), като фокусът е концептуализация на понятията за човешко тяло и тяхната класификация. Думи за означаване на глава, ръка и крак присъстват в лексиката на много езици.

* Валентина Стефанова. **Част на тялото** като семантична характеристика. – В: Св. Коева (ред.) *Семантични студии. Синтактично и семантично описание на български глаголи*. София: Издателство на БАН „Проф. Марин Дринов“, с. 46 – 73. <https://doi.org/10.7546/SemanticStudies2024.03>

Настоящото изследване¹ обръща перспективата. **Част на тялото (Body part)** се разглежда като семантична характеристика в структурата на глагола, където поне едно от съществителните имена, които глаголът налага, е изразено с фреймовия елемент **Част на тялото**. За описанието и анализа на тези глаголи се използват йерархично организирани семантични фреймове във Фреймнет според наличието на фреймов елемент **Част на тялото**. Съществителните имена, които глаголът присъединява, се разглеждат с оглед на отделни синонимни множества и семантични типове като възможна реализация на глагола. Целта е семантично описание и представяне на методология за извличане на информация и анализ на група глаголи според наличие на конкретен фреймов елемент, за да се постигне абстрактно описание на групи съществителни имена, които са подходящи за съчетаване при реализацията на глагола.

Настоящата разработка е част от създаването на Българския фреймнет (Булфрейм), в който за описание на българските лексикални единици се използват, ако е приложимо, семантичните фреймове от Фреймнет. За целта е въведена абстрактна структура (Коева 2021a), наречена „концептуален фрейм“, която може да съответства или не на езиково независима част от семантичните фреймове и в която към всеки фреймов елемент се описват семантичните класове на синонимни множества съществителни имена.

Изследването е свързано с описанието на множествата от съществителни имена, подходящи за реализация на фреймовия елемент **Част на тялото**, който е включен в описанието на различни семантични фреймове.

2. Ресурси

Използването на семантични ресурси, съотнасянето им спрямо релевантни признаци, обогатяването на лексикално-семантични мрежи и разработването на нови ресурси с помощта на Уърднет, Фреймнет, Върбнет и под. е дейност, която има дълга история и множество създадени ресурси за извличане на данни. Българският фреймнет (Коева / Коева 2010) се опира върху концептуалното знание, организирано в йерархичната структура на Фреймнет, а Българският уърднет (Коева 2021b) следва структурата на Принстънския уърднет (Уърднет), като през годините са правени множество съотнасяния на двете мрежи с цел обогатяване на българската лексикално-семантична мрежа.

Българският уърднет (Булнет) представлява лексикално-семантична мрежа, изградена по модела на Принстънския уърднет (Fellbaum 1998), в

¹ Първоначалните наблюдения по темата са представени по време на Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“ при Българската академия на науките през 2024 г. и публикувани в сборника от конференцията.

която езиковите единици са организирани в синонимни множества, свързани помежду си със семантични, морфо-семантични, деривационни и екстралингвистични релации (Коева 2021b). Лексикално-семантичната мрежа на Уърднет е съставена от взаимносвързани понятия и до голяма степен е езиково независима, което позволява създаването на семантични мрежи за различни езици, включително и за българския (Коева 2021b). Класификацията на синонимните множества в Уърднет (Fellbaum 1998), включваща „семантични класове“ или „семантични примитиви“, е пренесена от Принстънския уърднет и в Българския уърднет. Използвани са 25 класа за съществителните имена като: действия (noun.act), обекти, произведени от човека (noun.artifact), лица (noun.person), животни (noun.animal), растения (noun.plant), превозни средства (noun.vehicle), местоположения (noun.location), явления (noun.phenomenon) и други. Глаголите са разделени в 15 семантични класа, например глаголи за движение (verb.motion), за мисловна дейност и познавателни процеси (verb.cognition), за физическо състояние или действие по отношение на тялото (verb.body), за политически и социални дейности и събития (verb.social), за комуникация (verb.communication), за емоции (verb.emotions), за състояния (verb.stative) и други.

Фреймнет (Baker et al. 1998; Ruppenhofer et al. 2016) е езиков ресурс за английски, в който лексикалните единици са групирани в абстрактни семантични структури – семантични фреймове, отразяващи общи семантични и синтактични свойства. Различните значения на думите са свързани с различни семантични фреймове. Семантичните фреймове са снабдени с име и дефиниция и съдържат елементи на фрейма, които от своя страна също имат име, дефиниция, семантичен тип, спецификация за техния основен статус и вътрешни отношения между елементите на фрейма. Като цяло семантичните фреймове представляват схематични описания на концептуалната структура на ситуацията посредством участниците, обстоятелствата и други концептуални роли (елементи), които ги съставят (Fillmore et al. 2004: 1091).

Съотнасяне на семантични класове от Върбнет и семантични фреймове от Фреймнет със синонимни множества от Уърднет се предлага от Стоянова и Лесева (Stoyanova, Leseva 2023). Авторите разглеждат комбинацията от семантична и синтактична информация като възможен начин за трансфер на знания между езиците (напр. от английски на български), като се разчита на универсалността на семантичното описание. Като резултат са съотнесени семантични фреймове към синонимни множества в Уърднет, както и класове от Върбнет, представени в набора от 32 уникални семантични роли, съотнесени с общо 217 фреймови елемента от Фреймнет (Stoyanova, Leseva 2023).

Булнет е бил обект на обогатяване и с други семантични ресурси като семантичните типове от класификацията в Корпусния анализ на моделите (Corpus Pattern Analysis) (Hanks 2004), в резултат на което системата е допълнена със схематични глаголни модели и обобщени семантични типове на съществителните имена (Коева et al. 2018). Речникът на моделите на ан-

глийските глаголи (Patterns Dictionary of English Verbs) представя глаголни-те валентности, описани с помощта на семантичните типове от Корпусния анализ на моделите (КАМ). В същността си глаголните модели са типични употреби на определен глагол в контекст и представляват основната „аргументна структура“ на всеки глагол (със семантични стойности, посочени за всеки от елементите на моделите) (Hanks 2004: 87). Моделите се състоят от фиксиран йерархичен набор от семантични категории, чийто ред съответства на граматичните категории. Всеки модел е свързан с импликация, обясняваща „смисъла“ на модела, използвайки семантичните типове на КАМ (Hanks 2004: 88). Семантичните типове са формулирани посредством емпирично наблюдение на употребите в корпус и са подредени в сравнително плитка онтология.

Синонимните множества в Уърднет са обогатени чрез сливане на понятия от Уърднет и семантични типове от класификацията в Корпусния анализ на моделите, като хипоними на синонимни множества, към които е съотнесен семантичен тип от КАМ, наследяват не само съответния семантичен примитив на Уърднет, но и семантичния тип от КАМ. Съотнасянето се основава на хипотезата, че хипонимите наследяват семантичния тип на своите хипероними, и е осъществено чрез автоматична обработка и последваща ръчна проверка на данните (Коева et al. 2018). По този начин семантичната класификация на съществителните имена в Уърднет е разширена, за да се предостави по-точна информация за допустимостта или ограниченията при съчетаемостта между глаголи и съществителни имена (Коева et al. 2018). Например литералите eng-30-07891726-n²: {вино}, eng-30-07907943-n: {ликьор}, eng-30-07921615-n: {сайдер}, които са хипоними на синонимното множество eng-30-07884567-n: {алкохолно питие}, се характеризират не само със семантичния клас съществителни имена за храна, но получават и семантичен тип [Алкохолна напитка] [Alcoholic Drink].

Изреченските рамки, въведени в Уърднет, представляват плитка структура, описваща набор от аргументи, задължителни или незадължителни, които глаголът може да приеме, за да образува изречение (Някой (яде) Нещо; Някой (очаква) Някого и под.). След съотнасянето със семантичните типове от КАМ глаголите получават разширено представяне на аргументите, например: Човек | Животно яде Храна (Коева и др. / Коева et al. 2018: 69, 72).

Изследването, което правят Коева и колектив (Коева и др. / Коева et al. 2019), има за цел да конкретизира информацията за съчетаемостта на лексикалните единици. Поради тази причина е определен набор от лексикални единици, с които елементите от определен семантичен фрейм образуват цяло изречение. Стремещт е да се обогатят семантичните фреймове

² Синонимните множества се индексират с т. нар. междуезиков индекс (InterLingual Index – ILI), с помощта на който са свързани лексикалните съответствия в различните езици.

от Фреймнет с концептуална информация. На практика това означава да се уточни, че глаголет *дразня* например се комбинира с фреймов елемент **Експериенцер** (**Experiencer**), който може да бъде със семантичен тип [**Човек**] [**Human**], и с фреймов елемент **Стимул** (**Stimulus**), който може да бъде със семантичен тип [**Събитие**] [**Event**] или [**Същност**] [**Entity**] (Коева и др. / Коева et al. 2019: 552).

Трябва да се отбележи, че двата ресурса – Речникът на моделите на английските глаголи и Корпусният анализ на моделите – са неизменно свързани, семантичните типове описват аргументите в глаголният модел.

Изследването в настоящата студия се осъществява чрез извличане на информация от семантичните ресурси, създадени за целите на семантичното, синтактичното и концептуалното знание за глаголни единици. Както беше споменато във въведението, в основата на изследването стои и процесът по разработване на Булфрейм – система за визуализация и редакция на концептуални фреймове, които са дефинирани в системата (Коева, Doychev 2022).

Целта на концептуалната семантика според Джакендоф (Jackendoff 1992) може да се изрази чрез формулиране на няколко типа правила, които са тясно свързани и взаимнозависими: „а) правилата за формиране на концептуална структура, тоест примитиви (семантични класове) и принципи за комбинация, които заедно генерират безкраен клас от възможни концепции – включващи две понятия: лексикални (значение на думите) и фразови (изреченски понятия или съждения); б) прагматични и евристични правила; в) правилата за съответствие между концептуалната структура и други представяния, с които тя взаимодейства“ (Jackendoff 1992: 11).

Според Джакендоф (Jackendoff 1992) голяма част от теорията на концептуалната семантика се занимава с кодирането на предикатно-аргументната структура. Пример за концептуализация на определена синтактична структура дава чрез примери 1.а. и 1.б., където квадратните скоби в 1.б. идентифицират концептуалните съставни части на изречението.

1.а. [_S[_{NP} БИЛ]][_{VP}[_V ВЛЕЗЕ]][_{PP}[_P В]][_{NP} КЪЩАТА]]]

1.б. [_{СЪБИТИЕ} ([_{НЕЩО} БИЛ, [_{ПЪТ} КЪМ ([_{МЯСТО} В ([_{НЕЩО} КЪЩА))])])]]]

Джакендоф уточнява, че всяка съставна част от изречението е маркирана като принадлежаща към основна концептуална категория, или „семантична част на речта“, представляваща едно от понятията/концептите, съдържащи се в света, например Нещо (Thing) (или Физически обект (Physical object), Събитие (Event), Състояние (State), Път (Path) (или Траектория (Trajectory), Място (Place) (или Местоположение (Location), Време (Time), Сума (Amount) и под. (Jackendoff 1992: 13).

Коева (Коева 2021а) прави разграничение между семантичните фреймове във Фреймнет, съдържащи име на елемента, дефиниция, семантичен тип, и концептуалните фреймове, които съотнасят фреймовите елементи със синонимни множества на съществителни имена от Уърднет. Концептуалните фрей-

мове определят синтагматичните отношения на глаголите и съществителните. Те са абстрактни структури на семантичната и синтактичната съвместимост между глаголните предикати и съществителните имена като техни аргументи. Конкретен концептуален фрейм може да бъде свързан с определен семантичен тип от Фреймнет и всеки фреймов елемент от концептуалния фрейм е свързан с набор от съществителни, които са съвместими с предиката (Коева 2021а: 183). По-конкретно, концептуалният фрейм е семантичен фрейм по модела на Фреймнет, чиито ядрени фреймови елементи са асоциирани с множество от семантични класове от Уърднет (Коева 2021а: 183 – 184). Семантичните класове на Уърднет са обогатени със семантични типове според класификацията в Корпусния анализ на моделите и дефинират множества от лексикални единици в Уърднет, които удовлетворяват изискванията на ядрените елементи на глаголните фреймове във Фреймнет (Коева и др. / Koeva et al. 2019: 551).

Основната хипотеза, която се използва при разработването на системата Булфрейм, е, че всяко от синонимните множества в хиперонимно-хипонимната структура на Уърднет може да бъде представяне на лексикалната единица за тези елементи, като по този начин най-високото синонимно множество може да бъде избрано за описание на фреймовия елемент (Коева, Douchev 2022: 183). Например за фреймовия елемент **Продукти**, които се поглъщат (**Ingestibles**) на семантичния фрейм **Поглъщане (Ingestion)** съществително име, което се свързва с глагола *пия* (представител на посочения семантичен фрейм), може да бъде синонимното множество от Уърднет eng-30-07881800-n: {напитка; питие}, (noun, food), ‘всяка от разнообразни, най-често специално приготвени течности с определени вкусови качества, които удовлетворяват необходимостта на организма от прием на течности, имат освежаващо или друго действие или хранителна стойност’, на което е приписан семантичен тип **Напитка [Beverage]** и **Храна [Food]** от КАМ. Така {напитка; питие} и неговите хипоними BILI-00000069: {айрян}, BILI-00000007: {боза}, eng-30-07890970-n: {медовина}, eng-30-07844042-n: {прясно мляко}, eng-30-07921455-n {сайдер} могат да бъдат възможни аргументи на глагола *пия* (Коева, Douchev 2022: 210). Трябва да се има предвид обаче, че не всички синонимни множества в дадено семантично дърво биха били изцяло подходящи аргументи на определен предикат.

Наблюдението на реални езикови данни в корпус е от съществено значение за потвърждение на семантични ограничения. Лесева и Стоянова (Лесева, Стоянова / Leseva, Stoyanova 2020) посочват, че корпусните данни позволяват проверка на хипотезите за обхвата на семантичните ограничения, откриване на непоследователности, метафорични употреби и други особености на употребите в свободен текст. „Различните селективни ограничения за даден фреймов елемент може да се налагат и от глаголи, на които е приписан един и същи фрейм – например глаголи като *бръсна*, *сапунивам*, *правя маникюр* с фрейм **Грижа за тялото (Grooming)** изискват **Агентът (Agent)** да бъде лице, докато глаголи като *мия*, *почиствам* допускат както лица, така и животни. Изчерпателното описание на семантиката на глаголите включва и

обобщаване на селективните ограничения, които всеки от тях индивидуално налага“ (Лесева, Стоянова / Leseva, Stoyanova 2020: 243).

3. Изследвания, свързани с темата

Описването на глаголни синонимни множества от Уърднет чрез съпоставка с приписани семантични фреймове и анализирани фреймови елементи, които получават синтактична реализация като аргументи или адюнкти на глагола, е цел на изследване в отделни разработки, чийто обект са конкретни семантични класове, дефинирани от лексикално-семантичната мрежа.

Според Кукова (Кукова / Kukova 2023a) фреймовият елемент **Част на тялото** при глаголите за перцепция маркира сетивния орган, използван от активния или пасивния възприемащ при възприятието, който обикновено е изразен с предложна фраза (Кукова / Kukova 2023a: 33, Кукова / Kukova 2023b: 111), като повечето глаголи за възприятие съдържат имплицитно използвана **Част на тялото**. Например зрителното възприятие е задължително свързано с органите за зрение (Кукова / Kukova 2023a: 39 – 41).

В случаите, в които фреймовият елемент **Комуникатор (Communicator)** не е експлицитно изразен при глаголите за комуникация, може да се използва **Част на тялото** или действие с **Част на тялото** (поглед, кимване, жест и т. н.), с които да се предава желаната информация (Кукова / Kukova 2020: 236).

Според Лесева и Стоянова (Лесева, Стоянова / Leseva, Stoyanova 2021) при някои семантични класове, свързани с промяна, в които са включени глаголи за действие на тялото и неговото устройство и функциониране, може да се говори за стесняване, специализация на значението както при глагола {изчервявам се}, който представя промяна, свързана с (част на) тялото (Лесева, Стоянова / Leseva, Stoyanova 2021: 79).

Част на тялото показва какъв вид контактно действие се използва върху частта от тялото при глаголите за контакт, като специфичното при тези глаголи е, че в семантичния фрейм **Манипулация (Manipulation)** **Част на тялото** е конкретизирана във фреймовия елемент **Част на тялото на Агент (Bodypart of agent)** (Todorova 2023: 13 – 14).

Част на тялото се разглежда и като специфичен елемент за формулиране на семантични релации при Лесева и колектив (Лесева и др. / Leseva et al. 2018). Отделят се релации като Пациент.**Част на тялото** (част на тялото на участник, който търпи въздействие), Тема.**Част на тялото** (част на тялото, която е засегната от някакво въздействие), Локация.**Част на тялото** (част на тялото на участник в ситуацията (**Експериментер** и др.)), което е засегнато в хода на ситуацията) (Лесева и др. / Leseva et al. 2018: 23, 27, 34).

Някои от семантичните фреймове съдържат глаголи, които обикновено включват фреймов елемент **Част на тялото**, като при част от глаголите информация за този фреймов елемент се съдържа и в дефиницията на глагола. Например при глаголите за движение на тялото се очаква, че съществител-

ното, с което се изразява **Част на тялото**, ще бъде експлицитно изразено в изречението, както и че в дефиницията на глагола ще се включва конкретната **Част на тялото**: *куче маха с опашка* – ‘за човек или животно – правя движения с части от тялото си, обикн. крайник, глава и др.’ (РБЕ / РВЕ)³; *хората извиват вежди* ‘за коса, вежди, мустаци – падам, спускам се, завит на къдри или в дъгообразна форма’ (РБЕ / РВЕ) и т. н. В други случаи като *усмихвам се* ‘засмивам се леко и без глас’ (РБЕ / РВЕ) **Част на тялото** може да не бъде посочена в дефиницията. Тези случаи се приемат като инкорпорирана **Част на тялото**. Има случаи, при които **Част на тялото** може да бъде или да не бъде изразена в изречението, въпреки че е посочена в дефиницията на глагола (Ruppenhofer et al. 2016: 30): *тя мигна* или *тя мигна с очи* ‘често затварям и отварям очи’ (РБЕ / РВЕ).

Вежбицка (Wierzbicka 1988) представя три различни възможности за концептуализация на съществителните имена, които означават част на тялото от гледна точка на **Агент** и **Пациент**:

● **Част на тялото, разглеждана като отделен обект, която не е свързана с човека:**

2. *Питър изми неговата мръсна глава.*

(нещо се е случило с част Y от тялото на Z, защото X е направил нещо с нея)

3. *Питър си изми главата.*

(нещо се е случило с част Y от тялото на X, защото X е направил нещо с нея)

● **Част на тялото, разглеждана като обект, който е отделен, но свързан с човека:**

4. *Питър му изми неговата мръсна глава.*

(нещо се е случило с част Y от тялото на Z, защото X е направил нещо с нея; това, което се е случило с Y, може да се разглежда като нещо, което се е случило на Z)

5. *Питър си изми своята мръсна глава.*

(нещо се е случило с част Y от тялото на X, защото X е направил нещо с нея; това, което се е случило с Y, може да се разглежда като нещо, което се е случило с X)

● **Част на тялото, разглеждана като метонимична употреба на част от тялото:**

6. *Питър му изми мръсната глава.*

(нещо се е случило с част Y от тялото на Z, защото X е направил нещо с нея; мисля за него (за това, което се е случило с Y) като за нещо, което се е случило на Z, а не на нещо друго, различно от Z)

7. *Питър си изми мръсната глава.*

(нещо се е случило с част Y от тялото на X, защото X е направил нещо с нея; мисля за него (за това, което се е случило с Y) като за нещо, което

³ За осмисляне на значенията на съответния глагол са добавени дефиниции от *Речник на българския език* (онлайн) (РБЕ / РВЕ).

се е случило с X, а не с нещо друго, различно от X) (Wierzbicka 1988: 172 – 173).

Без да се навлиза в изследванията, засягащи въпроса за частите на тялото от анатомична гледна точка, както и тези, които сравняват наличието или липсата на конкретни понятия за част на тялото в различни езици, ще бъдат споменати няколко, които концептуализират част на тялото и свързаните с тялото понятия и представляват онтологичен интерес за настоящото изследване.

В описанието на Кант (Kant 1991)⁴ човешкото тяло се възприема като концептуализирано от гледна точка на три равнини (1. отгоре и отдолу, 2. двете страни на тялото (където са разположени ръцете) и 3. отпред и отзад): **Главата** се възприема като НАД всички други части на тялото, а **Краката** като ПОД всички останали части; **Ръцете** се възприемат като ОТ ДВЕТЕ СТРАНИ НА тялото; **Гърдите** и **Коремът** са разположени от ПРЕДНАТА страна на тялото, а **Гърбът** и **Седалището** са разположени от ЗАДНАТА страна на тялото.

Андерсен (Andersen 1978) предлага редица универсални правила за семантичната организация и категоризация на понятията, свързани с частите от човешкото тяло, които включват следното: 1) **Тялото** е обозначено във всички мероними на част от тялото; 2) Всеки език включва термин за **Глава**, който винаги принадлежи на **Тялото**; на второ ниво с **Глава** се срещат **Торс**, **Ръка** и **Крак**; 3) Всички езици имат понятия за **Очи**, **Нос** и **Уста**; 4) Горният крайник – **Ръка**, се означава по различен начин в различните езици (например в английски се различават понятията за *arm* (горната част на горния крайник) и *hand* (дисталната част на горния крайник, т. е. от китката до върха на пръстите), докато в български с **Ръка** се означават и двете английски понятия); 5) Повечето езици притежават понятие за английските термини *finger* и *toe* (съответно на български език **Пръст на ръка** и **Пръст на крак**), както и разграничени понятия за **Нокът на ръка** и **Нокът на крак** (*finger*)*nail* и (*toe*)*nail*); 6) Ако има отделен термин за **Крак** (*leg*), тогава ще има отделен термин и за **Ръка** (*arm*) (Andersen 1978: 352 – 353).

По-конкретно описание на частите от тялото представя Вежбицка (Wierzbicka 2006), която отделя пет основни части, съдържащи допълнителни съществителни имена за част на тялото: 1) **Крайници** (**ръка**, **рамо**, **крак**, **стъпало**); 2) **Глава** и свързаните с нея части (**глава**, **врат**, **очи**, **уши**, **нос**, **уста**, **език**); 3) **Лице** и „място по лицето“ (**чело**, **брадичка**, **бузи**); 4) **Торс** и неговите четири основни области (**гърди**, **корем**, **гърб**, **седалище**); 6) **Стави** (**китка**, **лакът**, **коляно**, **глезен**) (Wierzbicka 2006: 38 – 43).

4. Глаголи, свързани с човешкото тяло

Фреймовите елементи във Фреймнет се разделят на ядрени, периферни и екстратематични (Ruppenhofer et al. 2016: 19). Ядрените елементи предста-

⁴ Цитира се по Wierzbicka 2006.

вляват задължителни участници в семантичния фрейм, необходими за реализацията на дадено глаголно значение. В изследването се взима предвид конкретен фреймов елемент **Част на тялото**, спрямо който са извлечени глаголи, обект на изследването. Глаголите се наблюдават и анализират посредством наличните лексикално-семантични ресурси.

Това беше осъществено в няколко стъпки:

- 1) От Фреймнет бяха подбрани всички семантични фреймове⁵, в които един от ядрените фреймови елементи е **Част на тялото**, като тук се включва и фреймовият елемент **Част на тялото на Агент**, споменат в изследването на Тодорова (Todorova 2023).
- 2) Глаголите, които са описани с тези семантични фреймове⁶, бяха съотнесени с глаголни синонимни множества от Българския уърднет⁷. Първоначално съотнасянето беше осъществено автоматично. Поради многозначността на голям брой глаголи както в английски, така и в български като втора стъпка беше извършена ръчна проверка за потвърждение и подбор на литерали/глаголи спрямо дефиницията и семантичния клас на семантичните множества в Булнет. Въпреки наличните ресурси⁸, в които са приписани семантични фреймове на глаголи от Булнет, беше установено, че липсват семантични фреймове (с ядрен фреймов елемент **Част на тялото**), както и глаголи, които се описват чрез тях. Поради тази причина беше направено ново съотнасяне, като стимулът беше да се направи и допълнителна ръчна проверка.
- 3) От българското съотнасяне на синонимни множества в Булнет със семантични типове от КАМ⁹ бяха подбрани тези от тях, които отговарят на **Част на тялото** и неговите надредни и подредни множества спрямо онтологията на КАМ (виж. т. 5 в настоящата разработка).

Като резултат бяха идентифицирани 288 глагола от 11 семантични класа, описани в 15 семантични фрейма. Семантичните фреймове се разглеждат спрямо вътрешно йерархичната структура на Фреймнет чрез две релации: **Наследяване** и **Използване**.

Наследяване е най-силното отношение между семантичните фреймове. Релацията се осъществява между по-общ и по-конкретен фрейм, като по-конкретният детайлизира по-общия, а фреймовите елементи на по-конкретния съвпадат с елементите на по-общия или са тяхна конкретизация

⁵ <https://clres.com/clar/feindex.html>

⁶ <https://framenet.icsi.berkeley.edu/luIndex>

⁷ <https://dcl.bas.bg/bulnet/>

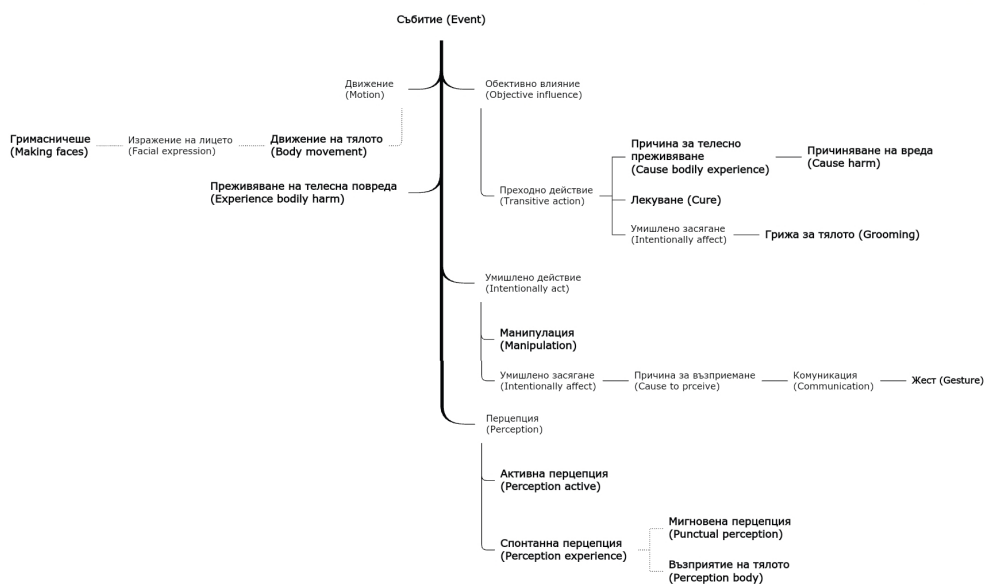
⁸ Семантичният ресурс: Автоматично приписани семантични фреймове от Фреймнет на глаголи от основния речников запас, е създаден през първия етап от изпълнението на проект „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“. <https://dcl.bas.bg/wp-content/uploads/2023/01/AutoAssignedFrames.pdf>

⁹ <https://dcl.bas.bg/PWN CPA/>

(Ruppenhofer et al. 2016: 75, Fillmore et al. 2004: 1092). Друга важна релация е **Използване**, която представя зависимост между по-общ и по-конкретен фрейм, като само част от елементите на по-общия имат съответствие в по-конкретния, като при това представляват тяхна конкретизация (Ruppenhofer et al. 2016: 78; Fillmore et al. 2004: 1092).

В структурата на Фреймнет има няколко семантични фрейма на най-високо абстрактно равнище¹⁰: **Събитие (Event)**, **Връзка (Relation)**, **Състояние (State)**, **Обект (Entity)**, **Място (Place)** и **Процес (Process)**. Семантичните фреймове са свързани с тях чрез релациите **Наследяване**, **Използване**, **Субфрейм** и др. Голяма част от семантичните фреймове се наследяват от **Събитие (Event)**.

На Фигура 1. се онагледява йерархичната свързаност между семантичните фреймове, които се разглеждат в настоящата работа и които включват ядрен фреймов елемент **Част на тялото**. Тези фреймове са показани в получер/удебелен стил. Непрекъснатите линии проследяват релацията **Наследяване**, а пунктирните линии – релацията **Използване**.



Фигура 1. Йерархична структура на семантични фреймове

4.1. Обективно влияние (Objective influence)

Семантичният фрейм **Обективно влияние** се дефинира¹¹ като „субект, ситуация или променлива оказват влияние върху друг субект, ситуация или променлива, които са поставени в зависима позиция“:

¹⁰ <https://framenet.icsi.berkeley.edu/FrameLatticeList>

¹¹ Използват се дефиниции на семантичните фреймове и фреймовите елементи, представени във Фреймнет.

4.1.1. **Причина за телесно преживяване (Cause bodily experience)** се дефинира като „Агент, който засяга **Част на тялото** на **Експериментатор** по такъв начин, че възприятието на **Експериментатора** се променя качествено“. Фреймовите елементи са **Агент**, **Експериментатор**, **Част на тялото**, **Причина (Cause)**, като **Част на тялото** представлява засегнато от **Агента** място върху тялото на **Експериментатора**. Идентифицирани са 7 глагола: за действие върху тялото (8.а.)¹² и за перцепция (8.б. и 8.в.).

8.а. {масажирам}¹³ ‘правя масаж на някого’ – *масажирам лицето, слепоочието, челото, ходилото, сухожилието, бедрото, стъпалото, кожата, гърба, пръста, крака, рамото* (синтактична функция – обект);

8.б. {чеша; разчесвам} ‘раздразвам наранено или възпалено, предизвикващо сърбене място чрез физически контакт с него’ – *чеша езика, носа, ухото, пресеца, брадата си* (с употреба на кратко възвратно притежателно местоимение, което показва, че **Част на тялото** принадлежи на субекта); *чеша я по главата, носа, врата, корема; зад ухото* (синтактична функция – обстоятелствено пояснение за място със задължителна кратка форма на личното местоимение); *до кръв* (синтактична функция – обстоятелствено пояснение за начин);

8.в. {гъделичкам} ‘докосвам част на тялото леко, така, че да възбудя повърхностните нерви и да предизвикам усещане за безпокойство, смях или спастични движения’ – *гъделичкам дланта, кожата, ухото, крака, пръста* (синтактична функция – обект); *по шията, корема; под мишницата, лъжичката* (синтактична функция – обстоятелствено пояснение за място); *с език* (синтактична функция – обстоятелствено пояснение за начин).

4.1.2 **Причиняване на вреда (Cause harm)** се дефинира като „ситуация, в която **Агент** наранява **Жертва (Victim)**, а **Част на тялото** на **Жертвата** е най-пряко засегната“. **Част на тялото** определя „мястото върху самото тяло, където е извършено телесното нараняване“. Фреймовите елементи са **Агент**, **Жертва**, **Част на тялото**, **Причина**. Идентифицирани са 67 глагола предимно от семантични класове за действие по отношение на тялото (9.а. и 9.б.) и за контакт (9.в. и 9.г.).

9.а. {ранявам} ‘причинявам физическа болка или увреждане’ – *ранявам коляното, слепоочието* (синтактична функция – обект); *в бедрото, рамото, корема, крака, гърба, сърцето, ръката, главата* (синтактична функция – обстоятелствено пояснение за място);

¹² Примерните употреби са събрани от системата Скеч Енджин (Sketch Engine) – софтуер за управление на корпуси и анализ на текст (<https://www.sketchengine.eu/>).

¹³ Глаголните единици се представят посредством синонимно множество от Българския уърднет (<https://dcl.bas.bg/bulnet/>).

9.б. {обезобразявам} ‘осакатявам жестоко, като засягам не само целостта и функциите, но и външния вид’ – *обезобразявам лицето, трупа, тялото* (синтактична функция – обект; предимно се означава по-голямо пространство от тялото);

9.в. {блъскам} ‘удрям, натискам или изтласквам някого или нещо с бутане’ – *блъскам главата, челото, тялото* (синтактична функция – обект); *в гърдите, по главата* (синтактична функция – обстоятелствено пояснение за място); *с лакът, юмрук, крак, ръка, глава* (синтактична функция – обстоятелствено пояснение за начин);

9.г. {бия} ‘нанасям удари с ръка’ – *бия главата (си/му)* (синтактична функция – обект; предпочита се употребата на кратко възвратно притежателно местоимение или кратко притежателно местоимение); *с ръка, юмрук* (синтактична функция – обстоятелствено пояснение за начин); *в главата; по главата, лицето, дупето; през пръстите, ръката, лицето, устата* (синтактична функция – обстоятелствено пояснение за място).

4.1.3 **Лекуване (Cure)** се дефинира като „**Агент**, който лекува и/или излекува наранявания, болести и причинена болка на **Пациент (Patient)**“. Фреймовите елементи са **Лечител (Healer)**, **Част на тялото**, **Пациент**, **Страдание (Affliction)**, **Лекарство (Medication)**, **Лечение (Treatment)**, като последните два се изразяват рядко. Идентифицирани са 9 глагола – глаголи, свързани с действия по отношение на човешкото тяло (10.а. и 10.б.) и за социални дейности и събития (10.в.).

10.а. {облекчавам} ‘подобрявам състоянието чрез намаляване на неприятно или болезнено (физическо) усещане’ – *облекчавам болката в стомаха, в гърлото, в ставите* (**Болка, Страдание** е задължителен елемент, а **Част на тялото** е конкретизация за мястото на болка, изразена с предложна фраза).

10.б. {заздравявам} ‘за орган или рана – оправям се след нараняване’ – *раната на главата, челото, бедрото, шията, коляното, крака, бузата, китката, ръката, веждата, врата, глезена, тила, корема, гърлото, рамото, пръста, крайника, ухото, устната, стъпалото, гърба, носа, клепача, петата заздравява* (съществителното име *рана* (и под.) е задължителен елемент и е в синтактична функция субект, докато **Част на тялото** е конкретизация за мястото на раната, изразена с предложна фраза в синтактична функция обстоятелствено пояснение за място); в примери като *клепачът заздравя, коляното заздравя, кожата заздравя* и др. **Част на тялото** е в синтактична функция субект. В подобни случаи *раната на клепача, разкъсването на кожата, изгарянето на кожата* и т. н. се приемат като вид метонимична употреба, както посочва и Вежбицка (Wierzbicka 1988).

10.в. {реhabilитирам} ‘възстановявам способността за адаптиране

към окръжаващата среда или към нови условия’ – най-често фреймов елемент **Част на тялото** не се изразява.

4.1.4. **Грижа за тялото (Grooming)** се дефинира като „Агент, който се ангажира с личната грижа за **Част на тялото** на Пациент чрез подстригване и поддръжка“. Фреймовите елементи са **Агент, Част на тялото, Пациент**. Идентифицирани са 17 глагола: глаголи за действия по отношение на тялото (11.а. и 11.б.), за контакт (11.в.) и за промяна (11.г.).

11.а. {бръсна} ‘отстранявам окосмението на **Част на тялото** с бръснач или подобно приспособление’ – *бръсна главата, брадата, мустака, мишницата, краката (си/му)* (синтактична функция – обект; предпочита се употребата на кратко възвратно притежателно местоимение или кратко притежателно местоимение); *бръсна (се/го) под мишницата* (синтактична функция – обстоятелствено пояснение за място; при реализация на **Част на тялото** с предложна фраза се предпочита употребата на кратко лично и възвратно лично местоимение);

11.б. {реша; сресвам; разресвам} ‘разглаждам и оформям (коса) с помощта на гребен или по подобен начин’ – *реша косата, козината, къдрицата, перуката* (синтактична функция – обект);

11.в. {сплитам} ‘правя на плитка кичур коса’ – *сплитам коса, кичур* (синтактична функция – обект);

11.г. {овлажнявам} ‘правя нещо влажно’ – *овлажнявам кожата, устната, косата* (синтактична функция – обект).

В тези случаи **Част на тялото** запълва позиция на пряк обект или непряк обект в предложна фраза и следва да бъде изразена в изречението. Ако действието е насочено към себе си, **Част на тялото** е пряк обект на глагола. При повечето глаголи за контакт от семантичния фрейм **Причиняване на вреда** **Част на тялото** може да се употреби с обстоятелствено пояснение за начин.

4.2. Умишлено действие (Intentionally act)

Като цяло **Умишлено действие** описва действия, които се извършват от съзнателни същества. В тази група спадат следните семантични фреймове:

4.2.1. **Манипулация** описва „акта на манипулиране на Обекта от страна на Субекта (**Агент**), като Обектът не е дълбоко или окончателно засегнат“. **Агентът** използва част от своето тяло за манипулиране на Обекта. Идентифицирани са 34 глагола от семантичен клас глаголи за контакт (12.а. и .12.б.). Фреймовите елементи са **Агент, Обект, Част на тялото на Агент**.

12.а. {хващам, държа} ‘улавям или поемам трайно с ръка (ръце) или в хватка’ – *хващам с ръка, с пръст (Част на тялото на Агента в предложна фраза); за гърлото, за ръката, за гушата (Част на тялото на Пациента в предложна фраза. Обособяване на човешко*

тяло може да се появи и с екстратематичен фреймов елемент **Локус (Locus)** или **Местоположение**, като се има предвид, че **Локус** може да обхваща и други понятия освен **Част на тялото**);

12.б. {опипвам} ‘докосвам някого грубо или по неприличен начин’ – *опипвам тяло, бюст* (синтактична функция – обект); *с пръст, с палец, с ръка, с крак* (синтактична функция – начин); *по гърба, по крака, по тялото* (синтактична функция – обстоятелствено пояснение за място).

От примерите се забелязва, че фреймовият елемент **Част на тялото** се изразява най-често като принадлежащ на засегнатия от действието на глагола или когато действието е насочено към себе си. Съществителните, изразяващи **Част на тялото** на **Агента**, обикновено се ограничават до *ръка* или *пръст* като обстоятелствено пояснение за начин на извършване на действието.

4.2.2. **Жест (Gesture)** описва „невербална комуникация. **Топикът** обикновено не се изразява с думи и може да се разглежда като частична характеристика на **Съобщението (Message)**, но такава частична характеристика е или трудна, или неуместна при липсата на вербален код“. **Част на тялото** е „специфична **Част на тялото** на **Комуникатора**“, чрез която се жестикулира, и е важно средство за осъществяване на комуникацията. Идентифицирани са 7 глагола за комуникация (13.а.). Фреймовите елементи са **Комуникатор**, **Адресат (Addressee)**, **Съобщение**, **Част на тялото**, **Идентифициран обект**.

13.а. {махвам} ‘давам знак чрез жест (с ръка)’; {сигнализирам} ‘съобщавам за нещо чрез определени знаци, сигнали’; {правя знак} ‘съобщавам информация или изразявам отношение, чувство, емоция и под. с помощта на движения с ръце, обикновено съпътстващи или заместващи словесното изразяване’ – *с ръка, с пръст, с глава*. Възможностите за изразяване на **Част на тялото** са ограничени до *ръка, пръст* и рядко *глава*.

4.3. Перцепция (Perception)

„**Възприемащият (Perceiver)** възприема определен **Феномен (Phenomenon)**. Семантичният фрейм е общ за всички наследяващи семантични фреймове, които описват действие на някакво съзнателно същество, реагиращо на промените в околната среда, независимо от вида на възприятието. Наследяващите семантични фреймове могат да уточняват видовете перцепция, да подчертават преживяванията или действията на възприемащия или свойствата на възприеманите явления.“

4.3.1. **Пасивна перцепция (Perception experience)** съдържа глаголи за перцепция, чрез които „**Възприемащият** получава знание от извършен опит, който обаче не е задължително да се възприема изцяло“. Обикновено се използват глаголи, които не изразяват специфично сетивно възприятие, като

откривам, възприемам, усещам. Идентифицирани са 11 глагола за перцепция. Фреймовите елементи са **Част на тялото, Пасивен възприемащ (Perceiver passive), Феномен**.

14.а. {чувам} ‘получавам информация или вести от или относно някого или нещо’ – *чух с ушите си, чуй ме с ушите си* (употребява се в мн. ч.; синтактична функция – начин);

14.б. {виждам} ‘виждам и осмислям образа’ – *виждам (неговото) лице* (синтактична функция – обект; предпочита се употребата на притежателно местоимение); *с очи* (синтактична функция – обстоятелствено пояснение за място).

4.3.2. **Активна перцепция (Perception active)** съдържа глаголи за перцепция, чрез които „**Възприемащият** умишлено насочва вниманието си към някаква същност или **Феномен**“. Фреймовите елементи са **Агентивен възприемащ (Perceiver agentive) и Феномен**. Трябва да се има предвид, че **Част на тялото** рядко се изразява, като по-скоро е имплицитно заложена в значението на глагола, отколкото експлицитно изразен в изречение. Идентифицирани са 24 глагола за перцепция (verb.perception) и, свързани с действия по отношение на човешкото тяло (verb.body). Например: *вкусвам (с уста), подслушвам (с уши), гледам (с очи)*.

Семантичните фреймове **Пасивна перцепция** и **Активна перцепция** са изследвани от Кукова (Кукова / Kukova 2023а; Кукова / Kukova 2023б).

4.3.3. **Мигновено възприятие (Punctual perception)** – „**Възприемащият** възприема за кратко **Феномен**, който често води до частично или несигурно възприятие. Фреймовият елемент **Част на тялото** може да се спомене като метонимичен заместител на **Възприемащия** или като част, която позволява самото възприятие“. **Част на тялото** е място върху тяло, където се осъществява възприятието, и обикновено се изразява с предложна фраза. Идентифициран е един глагол за перцепция (verb.perception).

15. {зървам} ‘хвърлям бърз поглед’ – *зърна лицето ѝ, зърна босия ѝ крак, зърна очите му*.

Примерите за български, които бяха намерени, показват, че фреймовият елемент **Част на тялото** може да се изразява и като принадлежащ на обекта, върху който пада действието. Фреймовият елемент определя зрителния орган, с който се извършва действието, и обикновено не се употребява експлицитно в изречението: *зърна (с очи) лицето ѝ*.

4.3.4. **Перцепция посредством тялото (Perception body)**

Перцепция посредством тялото се дефинира като семантичен фрейм, който описва „дадено физическо преживяване, засягащо част от тялото на човек“. Фреймовите елементи са **Експериенцер, Част на тялото**. Идентифицирани са 15 глагола, принадлежащи на няколко семантични класа.

16.а. {обгарям} ‘лекувам наранявания, образувания и други състояния или спирам кръвотечение чрез разрушаване на засегнатите тъкани с помощта на висока или много ниска температура, разяждащи химични вещества, електричество или радиация’ – *обгарям кожата* (синтактична функция – обект);

16.б. {гъделичкам} ‘докосвам леко някого или нещо, като предизвиквам лек, приятен сърбеж’ – *гъделичкам главата, ръката, тялото, гърба, пръста, петата* (синтактична функция – обект); *по пръста, ръката, главата, крака, гърба* (синтактична функция – обстоятелствено пояснение за място);

16.в. {сърби ме} ‘усещам дразнещо усещане по кожата’ – *сърби ме вратът, езикът, кожата, кракът, носът, ухото, тялото* (синтактична функция – субект);

16.г. {боде ме} ‘получавам усещане като от сърбеж или щипане’ – *боде ме в главата, ръката, крака, плешката* (синтактична функция – обстоятелствено пояснение за място).

4.4. Претърпяване на телесна вреда (Experience bodily harm)

Претърпяване на телесна вреда се дефинира като „**Експериенцер**, който търпи телесно нараняване на **Част на тялото**, която може и да не е експлицитно изразена в изречението“. **Част на тялото** специфицира мястото на нараняване върху **Експериенцера**. Фреймовите елементи са **Част на тялото**, **Експериенцер**. Идентифицирани са 19 глагола: глаголи, свързани с действия по отношение на тялото (17.а.), за контакт (17.б.) и за перцепция (17.в.).

17.а. {ранявам} ‘причинявам физическа болка или увреждане’ – *ранявам сърцето, тялото* (синтактична функция – обект); *в крака, сърцето, главата, ръката* (синтактична функция – обстоятелствено пояснение за място);

17.б. {разкъсвам} ‘прониквам в повърхност, като нарушавам целостта и здравината ѝ’ – *разкъсвам плътта, тъканта* (синтактична функция – обект); *със зъби, с нокти, с ръка* (синтактична функция – обстоятелствено пояснение за начин);

17.в. {навяхвам} ‘наранявам чрез удар, блъскане, дърпане и под., като разтягам или усуквам ставите’ – *навяхвам глезена, шията, ставата, коляното, крака, рамото* (синтактична функция – обект).

В повечето случаи **Част на тялото** се употребява като пряк обект в изречението, но при някои глаголи, както се вижда и от примерите, и като обстоятелствено пояснение за начин и място: {удрям} ‘рязко и със сила влизам в контакт и упражнявам физическо въздействие върху някого или нещо с част на тялото или с помощта на инструмент’ (verb.contact) – *удрям в/по главата, челото, лицето, ръката, гърба, тялото; в носа, корема, бъбрека, коляното, окото; под кръста* (обстоятелствено пояснение за място); *с глава, пръст* (обстоятелствено пояснение за начин).

4.5. Движение (Motion)

Движение се дефинира чрез „движението на определен обект, което започва от едно място **Източник (Source)** и завършва на друго място **Цел (Goal)**, като по време на движението е изминато пространство между двете точки **Път (Path)**“.

4.5.1. Движение на тялото (Body movement)

Движение на тялото се дефинира като семантичен фрейм, който „съдържа глаголи за движение или действие, което **Агентът** извършва, използвайки конкретна **Част на тялото** си или самото тяло“. **Агент** и **Част на тялото** са основните фреймови елементи, като при някои глаголи **Част на тялото** се „изразява от значението на глагола и може да не се употреби експлицитно в изречението“. Идентифицирани са 58 глагола: глаголи, свързани с действия по отношение на тялото (18.а.), глаголи за промяна (18.б.), за контакт (18.в.), за консумация (18.г.), за движение (18.д.).

18.а. {проточвам} ‘изпъвам силно (врата си), така че да мога да видя нещо или някого по-отблизо’ – *проточвам шия, гръбнак, мускул* (синтактична функция – обект);

18.б. {повалям се} ‘падам на земята в състояние на изтощение или смърт’ – самото тяло се поваля (синтактична функция – субект);

18.в. {сбръчквам} ‘събирам или свивам нещо, така че да се получат бръчки или гънки’ – *сбръчквам чело, носле, вежда, нос, устна* (синтактична функция – обект);

18.г. {мляскам} ‘отварям и затварям устата си бързо и шумно по време на хранене’ – *мляскам с уста, език* (синтактична функция – обстоятелствено пояснение за начин);

18.д. {поклащам} ‘движа напред-назад’ – *поклащам бедро, глава, ханш, корем, задник, крак, тяло, рамо* (синтактична функция – обект); *с глава* (синтактична функция – обстоятелствено пояснение за начин).

4.5.2. Гримасничене (Making faces)

Семантичният фрейм **Гримасничене** се дефинира по следния начин: „**Агентът** прави определено изражение на лицето, обикновено, но не задължително, когато причината е вътрешно емоционално състояние“. **Част на тялото** служи като заместител на **Агента**, като се посочва конкретна част от лицето на **Агента**. Фреймовите елементи са **Агент**, **Част на тялото**. Идентифицирани са 7 глагола за действие по отношение на човешкото тяло (19.а.) и за състояние (19.б.).

19.а. {всъя; смръщвам} ‘показвам неудоволствие с израза на лицето си’ – *смръщвам вежди, чело* (синтактична функция – обект);

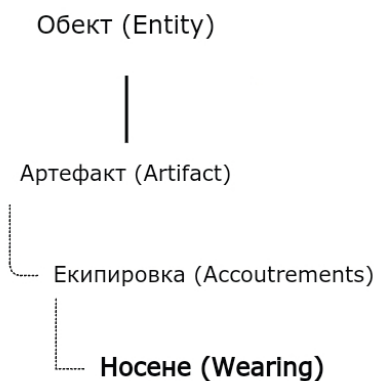
19.б. {мръщя се} ‘чувствам се сърдит или обиден и открито проявявам недоволството или раздражението си’ – *мръщя вежди, чело*

(синтактична функция – обект); *мръщя се с очи, уста, чело, вежди* (синтактична функция – обстоятелствено пояснение за начин).

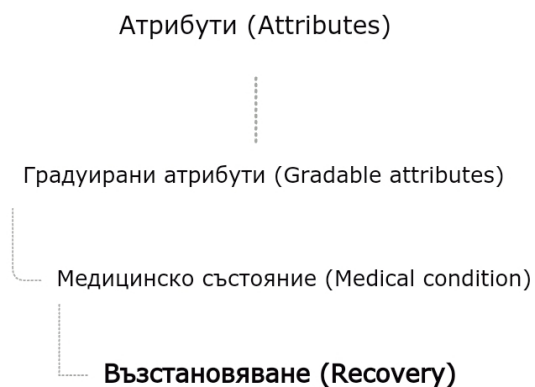
В повечето случаи **Част на тялото** не се изразява експлицитно. Ако бъде изразен като пряк обект или с предложна фраза, то тогава служи за подсилване значението на глагола, както при глагола *усмихвам се*: *усмихвам се от ухо до ухо, усмихвам се от сърце, усмихвам се с очи*.

4.6. Обект и Атрибути

Семантичният фрейм **Носене (Wearing)** е непряк наследник на **Обект** на Фигура 2., която обозначава различни обекти и свързаните с тях дейности. Семантичният фрейм **Възстановяване (Recovery)** е непряк наследник на **Атрибути** на Фигура 3. – по дефиниция изразява конкретен обект, който има определен атрибут с някаква стойност.



Фигура 2. Семантичен фрейм **Носене**



Фигура 3. Семантичен фрейм **Възстановяване**

4.6.1. Носене

Носене не се свързва с други глаголни семантични фреймове. Съдържа предимно прилагателни имена, но се срещат и три глагола – за контакт (20.а.), стативен глагол (20.б.) и глагол за действие по отношение на човешкото тяло (20.в.):

20.а. {разголвам} ‘откривам част на тялото напълно, особено такава, която обикновено е покрита’ – *разголвам гръд, бедро, телеса, плът, снага* (синтактична функция – обект);

20.б. {нося за показ} ‘обличам дреха върху тялото и демонстрирам гордост от това’ – *нося за показ върху тялото* (синтактична функция – обстоятелствено пояснение за място);

20.в. {нося} ‘облечен съм по определен начин’ – *нося върху тялото* (синтактична функция – обстоятелствено пояснение за място).

Дефинира се като семантичен фрейм, който описва „какво облекло има **Носещият (Wearer)**, като може да се посочи конкретна **Част на тялото на Носещия**“. Както се вижда от примерите, се предпочита или позоваване на самото тяло при глагола *нося*, или на по-малка част от него.

4.6.2. Възстановяване

Възстановяване се дефинира като семантичен фрейм, в който „глаголите описват процес на възстановяване или изцеление на **Пациент** от определена **Болка** без позоваване на конкретен процес на **Лечение** или **Лечител**, както при семантичния фрейм **Лечение**“. **Част на тялото** се дефинира като „отделна съставна част и се среща като външен аргумент“. За сравнение, при семантичния фрейм **Лечение** **Част на тялото** се среща като част от именна фраза, включваща съществителни като *рана*, *болка* и под. Идентифицирани са 6 глагола, като един е за движение (2.а.), а останалите са за действие по отношение на човешкото тяло (21.б.).

22.а. {отпускам} ‘придвижвам много бавно, щадящо и внимателно’ – *отпускам тяло, юмрук, ръка* (синтактична функция – обект);

22.б. {излекувам се} ‘оправям се от някакво заболяване или нараняване’; {възстановявам се} ‘възстановявам силите и енергията си след момент на спад и безсилие’ – *тялото се излекува / се възстанови* (синтактична функция – субект).

5. Селективни ограничения

Разглеждането на конкретни глаголи и анализът на семантични фреймове и фреймови елементи с тяхната синтактична реализация позволява по-прецизното определяне на възможните съществителни имена, които назовават фреймовия елемент **Част на тялото**. По този начин, от една страна, се осъществява конкретизация на значението на елемента, но от друга, се запазва концептуализацията на семантичните фреймове. Както беше посочено в т. 2, за тази цел се използва Корпусният анализ на моделите (Hanks 2004). В КАМ [Тяло][Body] е семантичен подтип на [Част от физически обект] [Physical Object Part]. От семантичния тип [Тяло] се отделя тип [Част на тялото] [Body Part], който се разделя на няколко допълнителни семантични подтипа: [Крайник] [Limb], [Сетивен орган] [Sense Organ], [Крак] [Foot], [Коса] [Hair], [Глава] [Head].

Обобщено, според данни след съотнасяне на лексикално-семантичната мрежа Уърднет със семантичните типове от КАМ (Koeva et al. 2018) посочените семантични типове, свързани с човешкото тяло, са приписани на 1857 синонимни множества.

Възможни семантични ограничения за фреймовия елемент **Част на тялото** на свързаните семантични фреймове с **Обективно влияние** в т. 4.1 са:

| | |
|---|---|
| <p>eng-30-05217168-n: {човешко тяло} [Тяло]: хипоним: eng-30-05219297-n: {юношеско тяло} хипоним: bg-eng-30-05219923-n: {женско тяло} хипоним: bg-eng-30-05219561-n: {тяло на възрастни} хипоним: eng-30-05219724-n: {мъжко тяло}</p> | <p>eng-30-05225090-n: {външна част на тялото} [Част на тялото]: хипоним: eng-30-01514549-n: {гърло} хипоним: eng-30-05305806-n: {уста} хипоним: eng-30-05538625-n: {глава} хипоним: eng-30-05546540-n: {врат; шия} хипоним: eng-30-05601198-n: {лицева част} хипоним: eng-30-01514431-n: {гърди} и др.</p> |
| <p>eng-30-05560244-n: {крайник} [Крайник]: хипоним: eng-30-02464461-n: {заден крайник} хипоним: eng-30-02464785-n: {преден крайник} хипоним: eng-30-05561390-n: {подбедрица} хипоним: eng-30-05561507-n: {крак} хипоним: eng-30-05562249-n: {бедро} хипоним: eng-30-05563770-n: {ръка} Хипоним: BUL-1889038884: {долен крайник}</p> | <p>eng-30-05254795-n: {космена покривка} [Коса]: хипоним: eng-30-01898731-n: {козина}; хипоним: eng-30-01899746-n: {грива} хипоним: eng-30-02416270-n: {брада} хипоним: eng-30-05256862-n: {прическа} хипоним: eng-30-05257737-n: {кичур; къдрица} хипоним: eng-30-05261404-n: {лицево окосяване} хипоним: eng-30-05261566-n: {брада} хипоним: eng-30-05262185-n: {мустак} хипоним: eng-30-05313535-n: {вежда} хипоним: eng-30-01900006-n: {перчем} хипоним: eng-30-05255578-n: {телесно окосяване}</p> |

Таблица 1. Семантични ограничения за фреймовия елемент **Част на тялото**

В тази група към **Част на тялото** се причисляват както цялото тяло, така и негови отделни части, които могат да бъдат засегнати от действието на глагола. Тук спадат голяма част от синонимните множества поради разнообразния обхват на глаголите. Така например за семантичните фреймове **Причина за телесно преживяване** и **Причиняване на вреда** съществителните имена, които глаголът присъединява, се отнасят до семантичното множество на [Тяло], [Част на тялото] и/или [Крайник]. От друга страна, в семантичния фрейм **Грижа за тялото** фреймовият елемент **Част на тялото** се изразява предимно чрез окосмената част на тялото: eng-30-05254795-n: {космена покривка} [Коса], докато в семантичния фрейм **Лекуване** може да се изрази като eng-30-05225090-n: {външна част на тялото} [Част на тялото], но и eng-30-05298729-n: {вътрешен орган} [Част на тялото].

При глаголите от т. 4.2 съществителните имена, които означават фреймовия елемент **Част на тялото** на семантичния фрейм **Умишлено действие**, се свеждат до eng-30-05563770-n: {ръка} [Крайник]; eng-30-05566097-n: {пръст} [Част на тялото] или eng-30-05538625-n: {глава} [Глава] поради семантиката на идентифицираните глаголи.

Възможни семантични ограничения за фреймовите елементи на семантичния фрейм **Перцепция** са:

- eng-30-05299178-n: {сетивен орган} [Сетивен орган];
- хипоним – eng-30-05300231-n: {хеморецептор};
- хипоним – eng-30-05598147-n: {нос};
- хипоним – eng-30-05311054-n: {око};
- хипоним – eng-30-05320899-n: {ухо}.

Предимно се избират сетивни органи, чрез които се възприема зрителино, слухово или обонятелно. В някои случаи се избират и глаголи за тактилно възприятие, при които има контакт с кожата. Възможните съществителни имена се свеждат до eng-30-05238282-n: {кожа} [Част на тялото] или eng-30-05225090-n: {външна част на тялото} [Част на тялото]. **Претърпяване на телесна вреда** и **Движение на тялото** подобно на семантичните фреймове **Причина за телесно преживяване**, **Причиняване на вреда**, **Лекуване** приемат множество съществителни имена и най-вече хипоними на семантични множества като: eng-30-05217168-n: {човешко тяло} [Тяло], eng-30-05225090-n: {външна част на тялото} [Част на тялото], eng-30-05560244-n: {крайник} [Крайник].

Възможните съществителни имена, които глаголите присъединяват в семантичния фрейм **Гримасничене**, се свеждат до изрази на човешкото лице, съответно: eng-30-05600637-n: {лице} [Част на тялото]; eng-30-05602548-n: {чело} [Част на тялото]; eng-30-05313535-n: {вежда} [Коса]; до някаква степен и eng-30-05301908-n: {уста} [Част на тялото].

В Таблица 2. е направено разпределение на съществителните имена, които глаголът налага, основано на класификацията на частите на тялото по Вежбицка (Wierzbicka 2006), на които са приписани синонимни множества от Българския уърднет и асоциираните семантични типове по КАМ. В резултат от изследването към класовете на Вежбицка са добавени нови съществителни имена за изразяване на **Част на тялото**, представени в последната колона на таблицата, както и подбрани хипонимни синонимни множества от Булнет.

| | | |
|--|--|--|
| eng-30-05560244-n: {крайник} [Крайник] | eng-30-05563770-n: {ръка} [Крайник] | eng-30-05566097-n: {пръст} [Част на тялото] |
| | | eng-30-05581693-n: {ноктъ} [Част на тялото] |
| | | eng-30-05565696-n: {длан} [Част на тялото] |
| | eng-30-05548840-n: {рамо} [Част на тялото] | - |
| | eng-30-05561507-n: {крак} [Крайник] | eng-30-05566097-n: {пръст} [Част на тялото] |
| | | {ноктъ} [Част на тялото] |
| eng-30-05562249-n: {бедро} [Крайник] | | |
| eng-30-05563266-n: {стъпало} [Крак] | eng-30-05578095-n: {пета} [Част на тялото] | |
| | - | |
| eng-30-05538625-n: {глава} [Глава] | eng-30-05538625-n: {глава} [Глава] | eng-30-01514549-n: {гърло} [Част на тялото] |
| | eng-30-05546540-n: {врат; шия} [Част на тялото] | eng-30-05602982-n: {гуша} [Част на тялото] |
| | eng-30-05311054-n: {око} [Сетивен орган] | - |
| | eng-30-05598147-n: {нос} [Сетивен орган] | - |
| | eng-30-05301908-n: {уста} [Част на тялото] | - |
| | eng-30-05301072-n: {език} [Част на тялото] | - |
| | eng-30-05320899-n: {ухо} [Сетивен орган] | - |
| eng-30-05600637-n: {лице} [Част на тялото] | eng-30-05602548-n: {чело} [Част на тялото] | - |
| | eng-30-05599617-n: {брадичка} [Част на тялото] | - |
| | eng-30-05602835-n: {буза} [Част на тялото] | - |

| | | |
|--|--|--|
| eng-30-05549830-n: {торс} [Част на тялото] | eng-30-01514431-n: {гърди} [Част на тялото] | – |
| | eng-30-05556943-n: {корем} [Част на тялото] | – |
| | eng-30-05558717-n: {гръб} [Част на тялото] | eng-30-05588174-n: {гръбнак} [Част на тялото] |
| | eng-30-05559256-n: {седалище} [Част на тялото] | – |
| eng-30-05595083-n: {става} [Част на тялото] | eng-30-05584928-n: {китка} [Част на тялото] | – |
| | eng-30-05579944-n: {лакът} [Част на тялото] | – |
| | eng-30-05573602-n: {коляно} [Част на тялото] | – |
| | eng-30-05578442-n: {глезен} [Част на тялото] | – |
| eng-30-05254795-n: {космена покривка} [Коса] | eng-30-05256085-n: {коса} [Коса] | – |
| | eng-30-02416270-n: {брада} [Коса] | – |
| | eng-30-05313535-n: {вежда} [Коса] | – |
| | eng-30-05261404-n: {лицево окосмяване} [Коса] | – |
| | eng-30-05255578-n: {телесно окосмяване} [Коса] | – |
| eng-30-05269901-n: {кост}[Част на тялото] | eng-30-05540121-n: {череп} [Част на тялото] | – |
| | eng-30-05546040-n: {челюст} [Част на тялото] | – |
| | eng-30-05279026-n: {ребро} [Част на тялото] | – |
| | eng-30-05284333-n: {прешлен} [Част на тялото] | – |
| | eng-30-05282746-n: {зъб} [Част на тялото] | – |
| | eng-30-05274247-n: {опашна кост} [Част на тялото] | – |
| | eng-30-05278584-n: {фаланга} [Част на тялото] | – |

Таблица 2. Класификация на Част на тялото

6. Заключение

В множество случаи фреймовият елемент **Част на тялото** запълва позиция на пряк обект, като при някои от тях употребата на кратка форма на притежателното местоимение, възвратното притежателно и възвратното лично местоимение е задължителна. Изразяването на **Част на тялото** като обстоятелствено пояснение се употребява най-често за място и начин.

От примерите се забелязва, че **Част на тялото** се изразява най-често като принадлежаща на засегнатия от действието на глагола или когато действието е насочено към себе си. **Част на тялото** може да се изрази и метафорично, за подсилване на значението на глагола: *усмихвам се с очи*.

Обобщено, **Част на тялото** се явява като фреймов елемент, който без конкретни примери може да се изразява с всяка една дума за обозначаване на **Част на тялото**. След внимателен преглед на глаголите може да се предложат по-конкретна класификация на съществителните имена, асоциирани чрез семантични типове и спрямо тяхната синтактична реализация. Допълнителен фактор за подобно определяне е и семантичният фрейм, който специфицира употребата на глагола в определен контекст. Както се вижда и от предложеното изследване, подобно ограничаване е релевантно например за глаголи от семантичните фреймове за перцепция, при които перцепцията се изразява чрез конкретни сетивни органи. Подобно е отношението и при глаголите от семантичния фрейм **Грижа за тялото**, където съществителните, изразяващи част на тялото, се ограничават до означаване на космената покривка на тялото. Глаголите за изразяване на **Жест** също се ограничават до определена **Част на тялото** – ръка, пръст и в редки случаи глава, както и при семантичния фрейм **Гримасничене**, където употребата на съществителни имена за **Част на тялото** се свежда до човешкото лице: eng-30-05600637-n: {лице} [Част на тялото]; eng-30-05602548-n: {чело} [Част на тялото]; eng-30-05301908-n: {уста} [Част на тялото]; eng-30-05313535-n: {вежда} [Коса].

Благодарности: Изследването е извършено по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания“, Договор № КП-06-Н 50/1 от 30.11.2020 г.

Цитирана литература / References

- Коева 2010: Коева, С. *Българският ФреймНет*. София: Институт за български език „Проф. Любомир Андрейчин“. (Koeva 2010: Koeva, S. *Balgarskiyat FreymNet*. Sofia: Institut za balgarski ezik „Prof. Lyubomir Andreychin“.)
- Коева и др. 2019: Коева, С., Т. Dimitrova, V. Stefanova, D. Hristov. Towards Conceptual Frames. – В: Чуждозезиково обучение, 46, 6, 2019, с. 551 – 564. (Koeva et al. 2019: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Towards Conceptual Frames. – In: Chuzhdoezikovo obuchenie, 46, 6, 2019, с. 551 – 564.)

- Кукова 2020: Кукова, Х. Глаголи за комуникация, фреймови елементи и семантични ограничения (върху синонимни множества от Булнет). – В: Мичева, В., Д. Благоева, М. Витанова, М. Цибранска, С. Колковска, Т. Александрова (ред.). *Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“* (София, 2020). Т. 2. София: Издателство на БАН „Проф. Марин Дринов“, с. 233 – 241. (Kukova 2020: Kukova, H. Glagoli za komunikatsiya, freymovi elementi i semantichni ograniicheniya (varhu sinonimni mnozhestva ot Bulnet). – In: Micheva, V., D. Blagoeva, M. Vitanova, M. Tsibranska, S. Kolkovska, T. Aleksandrova (Eds.). *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik “Prof. Lyubomir Andreychin”* (Sofia, 2020). T. 2. Sofia: Izdatelstvo na BAN “Prof. Marin Drinov”, pp. 233 – 241.)
- Кукова 2023а: Кукова, Х. Предикати за активна перцепция и особености на фреймовите им елементи. – *Известия на Института за български език „Проф. Любомир Андрейчин“*, XXXVI, с. 29 – 47. (Kukova 2023a: Kukova, H. Predikati za aktivna pertseptsiya i osobenosti na freymovite im elementi. – *Izvestiya na Instituta za balgarski ezik “Prof. Lyubomir Andreychin”*, XXXVI, pp. 29 – 47.)
- Кукова 2023б: Кукова, Х. Глаголи за спонтанна перцепция в българския език. – В: *Материали от Третата международна славистична конференция, посветена на паметта на св. св. Кирил и Методий*. Киев – София – Харков – Шумен: Харковско историко-филологическо дружество, с. 108 – 113. (Kukova 2023b: Kukova, H. Glagoli za spontanna pertseptsiya v balgarskiya ezik. – In: *Materiali ot Tretata mezhdunarodna slavistichna konferentsiya, posvetena na pametta na sv. sv. Kiril i Metodiy*. Kiev – Sofia – Harkov – Shumen: Harkovsko istoriko-filologicheskoto druzhestvo, pp. 108 – 113.)
- Лесева и др. 2018: Лесева, Св., Ив. Стоянова, Хр. Кукова, М. Тодорова. Интегриране на субкатегоризационна информация в релационната структура на УърдНет. – *Български език*, 65, 2, с. 11 – 40. (Leseva et al. 2018: Leseva, Sv., Iv. Stoyanova, Hr. Kukova, M. Todorova. Integrirane na subkategorizatsionna informatsiya v relatsionnata struktura na WordNet. – *Balgarski ezik*, 65, 2, pp. 11 – 40.)
- Лесева, Стоянова 2020: Лесева, Св., Ив. Стоянова. Отвѣд лексикалните ресурси: валидиране на концептуалното описание в корпусни данни. – В: *Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“* (София, 2020 г.). Т. 2. София: Издателство на БАН „Проф. „Марин Дринов“, с. 241 – 249. (Leseva, Stoyanova 2020: Leseva, Sv., Iv. Stoyanova. Otvd leksikalnite resursi: validirane na kontseptualnoto opisane v korpusni dannii. – In: *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik “Prof. Lyubomir Andreychin”* (Sofia, 2020 g.). T. 2. Sofia: Izdatelstvo na BAN “Prof. Marin Drinov”, pp. 241 – 249.)
- Лесева, Стоянова 2021: Лесева, Св., Ив. Стоянова. Семантично описание на глаголи за промяна и йерархична организация на концептуалните фреймове. – В: *Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“*. Т. 2. София: Издателство на БАН „Проф. Марин Дринов“, с. 76 – 85. (Leseva, Stoyanova 2021: Leseva, Sv., Iv. Stoyanova. Semantichno opisane na glagoli za promyana i yerarhichna organizatsiya na kontseptualnite freymove. – In: *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik “Prof. Lyubomir Andreychin”*. T. 2. Sofia: Izdatelstvo na BAN “Prof. Marin Drinov”, pp. 76 – 85.)

- РБЕ: *Речник на българския език (онлайн). А – системност*. <<https://ibl.bas.bg/rbe/>> [20.03.2024] (RBE: *Rechnik na balgarskiya ezik (online). A – sistemnost*. <<https://ibl.bas.bg/rbe/>> [20.03.2024])
- Andersen 1978: Andersen, E. S. Lexical universals of body part terminology. – In: Greenberg, J. H. (Ed.). *Universals of Human Language. Vol. 3. Word Structure*. Stanford: Stanford University Press, pp. 335 – 368.
- Baker et al. 1998: Baker, C., C. Fillmore, J. Lowe. The Berkeley FrameNet Project. – In: *COLING-ACL'98: Proceedings of the Conference*. Montreal, Canada, pp. 86 – 90.
- Brown 1976: Brown, C. H. General principles of human anatomical paratomy and speculations on the growth of paratonomic nomenclature. – *American Ethnologist*, vol. 3, no. 3, Folk Biology, pp. 400 – 424.
- Enfield 2006: Enfield, N. J. Elicitation guide on parts of the body. – *Language Sciences*, vol. 28, no. 2 – 3, pp. 148 – 157. <<https://doi.org/10.1016/j.langsci.2005.11.003>> [28.03.2024]
- Enfield et al. 2006: Enfield, N. J., A. Majid, M. van Staden. Cross-linguistic categorization of the body: Introduction. – *Language Sciences*, vol. 28, 2 – 3, pp. 137 – 147. <<https://doi.org/10.1016/j.langsci.2005.11.001>> [28.03.2024]
- Fellbaum 1998: Fellbaum, C. (Ed.). *WordNet: An Electronic Lexical Database*. Cambridge (MA): MIT Press.
- Fillmore et al. 2004: Fillmore, Ch. J., C. F. Baker, H. Sato. FrameNet as a “Net”. – In: *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC'04)*. Lisbon: European Language Resources Association (ELRA). <<http://www.lrec-conf.org/proceedings/lrec2004/pdf/388.pdf>> [24.02.2024]
- Hanks 2004: Hanks, P. Corpus pattern analysis. – In: *Proceedings of the 11th EURALEX International Congress (EURALEX 2004)*. Lorient: Universit  Bretagne Sud, pp. 87 – 97.
- Jackendoff 1992: Jackendoff, R. Parts and boundaries. – In: Levin, B., P. Steven (Eds.). *Lexical and conceptual semantics*. Cambridge, Ma.: Blackwell, pp. 9 – 45.
- Kant 1991: Kant, I. Von dem Ersten Grunde des Unterschiedes der Gegenden im Raume (Translation ‘On the First Ground of the Distinction of Regions in Space’). – In: van Cleve, J., R. E. Frederick (Eds.). *The Philosophy of Right and Left*. Dordrecht: Kluwer, pp. 27 – 33.
- Koeva 2021a: Koeva, S. Towards Expanding WordNet with Conceptual Frames. – In: Bosch, S., Ch. Fellbaum, M. Griesel, A. Rademaker, P. Vossen (Eds.). *Proceedings of the 11th Global Wordnet Conference*. Potchefstroom, South Africa, pp. 181 – 191.
- Koeva 2021b: Koeva, S. The Bulgarian WordNet: Structure and specific features. – In: *Papers of BAS*, 8, 1, pp. 47 – 70.
- Koeva et al. 2018: Koeva, S., T. Dimitrova, V. Stefanova, D. Hristov. Mapping WordNet concepts with CPA ontology. – In: *Proceedings of the 9th Global WordNet Conference (GWC'2018)*. Singapore: Global WordNet Association, pp. 70 – 77.
- Koeva, Doychev 2022: Koeva, S., E. Doychev. Ontology supported frame classification. – In: *Proceedings of the Fifth International Conference on Computational Linguistics in Bulgaria* Sofia: Department of Computational Linguistics, IBL – BAS, pp. 203 – 214. <<https://aclanthology.org/2022.clib-1.23>> [24.02.2024]
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: Extended Theory and Practice (Revised November 1, 2016)*. <<https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>> [24.02.2024]

- Stoyanova, Leseva 2023: Stoyanova, I., S. Leseva. Expanding the Conceptual Description of Verbs in WordNet with Semantic and Syntactic Information. – In: *Proceedings of the Global WordNet Conference (GWC 2023)*. Donostia: Global Wordnet Association, pp. 284 – 294.
- Todorova 2023: Todorova, M. A. Semantic Annotation of Verbs of Contact in Bulgarian. – In: *Proceedings of the 19th Joint ACL-ISO Workshop on Interoperable Semantics (ISA-19)*. Nancy: Association for Computational Linguistics, pp. 11 – 17.
- Wierzbicka 1988: Wierzbicka, A. *Studies in Language Companion Series, 18. The semantics of grammar*. Amsterdam & Philadelphia: John Benjamins. <doi.org/10.1075/slcs.18> [30.03.2024]
- Wierzbicka 2006: Wierzbicka, A. Bodies and their parts: An NSM approach to semantic typology. – *Language Sciences*, 29 (2007), pp. 14 – 65. <doi.org/10.1016/j.langsci.2006.07.002> [30.03.2024]

BODY PART AS A SEMANTIC FEATURE

Valentina Stefanova

Institute for Bulgarian Language „Prof. L. Andreychin“

valentina@dcl.bas.bg

Abstract. This study focuses on the analysis of the frame element **Body part** as a semantic feature to describe semantic frames according to the hierarchical structure in FrameNet. The aim is a semantical description and a methodological presentation for information extraction and analysis of a group of verbs according to the presence of a specific frame element. The research is part of the development of the Bulgarian FrameNet (BulFrame), a system for visualisation and editing of conceptual frames that are defined in the system (Koeva, Doychev 2022). The analysis is carried out by extracting verbs that are conceptually described by semantic frames with the core frame element **Body Part**. Semantic resources that are created for the purpose of semantic, syntactic and conceptual knowledge of verbs are used in the study. Mapping with verb synsets from the Bulgarian WordNet is made for the observation. The nouns that the verb conjuncts are considered with respect to distinct synsets and semantic types as a possible realisation of the verb. As a result, a grouping of nouns associated with semantic types from Corpus Pattern Analysis (CPA) is proposed.

Keywords: *semantic frame, frame element, semantic type, body part*

Valentina Stefanova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria

<https://orcid.org/0000-0003-3334-2050>

VERBS OF COMMUNICATION: VALIDATION OF SEMANTIC FRAMES IN CORPUS DATA

Svetlozara Leseva, Ivelina Stoyanova

Institute for Bulgarian Language, Bulgarian Academy of Sciences

zarka@dcl.bas.bg, iva@dcl.bas.bg

Abstract. The chapter outlines the properties of the semantic class of verbs of communication in terms of the most representative FrameNet frames of higher frequency and the syntactic realisation of the frame elements in different valence patterns in English and Bulgarian. For the purposes of the study we employ two large lexical-semantic resources: (a) the Princeton WordNet (Fellbaum 1998b), and the Bulgarian WordNet (Koeva 2021), and (b) FrameNet (Baker et al. 1998). In particular, the analysis is centred on the information included in each of the resources and how it can be used towards their mutual enrichment and the extension of their coverage.

We discuss the general organisation of the verb lexis representing the domain of communication: the prototypical frame **Communication** and frames inheriting from it, focusing on the frames **Statement** and **Telling** as two of the most representative frames of verbal communication. The objective is to validate the realisation of semantic frames in corpus data using the semantically annotated corpora SemCor (Miller et al. 1993b) and BulSemCor (Koeva et al. 2006). While we use resources for English and Bulgarian, we discuss the universal and language-specific aspects of this description and the transferability of knowledge across languages.

The observations made on the valence patterns and the syntactic expression of the core frame elements are used to verify the validity of the assigned frame, while also highlighting the similarities and differences both between verbs from the same domain in one language (Bulgarian) and between equivalent/similar senses across languages (Bulgarian and English).

Keywords: *verbs of communication, WordNet, FrameNet, valence patterns, corpus*

* Svetlozara Leseva, Ivelina Stoyanova. Verbs of Communication: Validation of Semantic Frames in Corpus Data. – In: Koeva, Sv. (Ed.) *Semantic studies. Syntactic and semantic description of Bulgarian verbs*. Sofia: BAS Press „Prof. Marin Drinov“, pp. 74 – 116. <https://doi.org/10.7546/SemanticStudies2024.04>

1. Motivation

We focus on the complex semantic description of verbs as represented in lexical semantic resources (such as WordNet and FrameNet) containing complementary semantic and valence information. We discuss both the language-independent and the language-specific aspects of conceptual description with a view to the transferability of knowledge between the two languages. We illustrate our findings through a case study on verbs of communication in English and Bulgarian. Further, we compile a corpus of illustrative examples in English and Bulgarian in order to observe the semantic and syntactic properties determining the realisation of each verb and the core frame elements in its evoked frame.

The class of verbs of communication comprises a diverse set of verbs with high frequency in the language, thus particularly suitable to illustrate the relations of inheritance between the prototypical frame **Communication** and the frames that inherit from it, such as **Communication manner**, **Statement**, **Telling**, etc. Moreover, the verbs belonging to the domain of communication are characterised by typical frame elements exhibiting certain semantic properties, e.g. the **Communicator** and the **Addressee** are sentient beings able to take part in communication. The frame elements are realised in typical syntactic positions. The observations on the corpus data aim to identify any configurations for the realisation of the verbs in English and Bulgarian that are confirmed for both languages or are valid for only one of them.

For the purposes of the study we employ two main lexical-semantic resources: (a) the Princeton WordNet (PWN) (Fellbaum 1998b), and the Bulgarian WordNet (Koeva 2021), and (b) FrameNet (Baker et al. 1998; Ruppenhofer et al. 2016). Illustrative examples are extracted from various sources: (i) annotated examples on English verbs from FrameNet; (ii) examples from the semantically annotated corpora SemCor (Miller et al. 1993b) for English and BulSemCor (Koeva et al. 2006) for Bulgarian; (iii) parallel examples from the Bulgarian-English Sentence- and Clause-Aligned Corpus (BulEnAC) (Koeva et al. 2012a).

The chapter is organised as follows. Section 2 presents the resources employed in the study with a focus on lexical-semantic resources (WordNet and FrameNet) and the corpus data from which illustrative examples are extracted and annotated. Section 3 provides an overview of related works focusing on other similar lexical-semantic resources and their relevance for the study. Section 4 describes the principle of universality of semantic features in conceptual resources. Section 5 presents the domain of communication and the frames that are used to describe aspects of it, focusing on the prototypical, “basic” frame **Communication**, and some of the most representative ones – **Statement** and **Telling**. More detailed observations on Bulgarian verbs from the communication domain are offered in Sections 6 and 7. The final section draws conclusions and discusses directions for future work.

2. Resources

Below we describe the lexical semantic resources and the corpora used to obtain data for the study.

2.1. Lexical-semantic resources

WordNet¹ (Miller et al. 1993a; Miller 1995; Fellbaum 1998b) is a large lexical database that represents comprehensively conceptual and lexical knowledge in the form of a network whose nodes denote synonym sets (synsets) linked by means of a number of semantic and lexical relations such as hypernymy, meronymy, antonymy, etc. We use both the Princeton WordNet and the Bulgarian WordNet (Koeva 2014; Koeva 2021), which are aligned at the synset level by means of unique synset identifiers.

In WordNet a coarsely-grained semantic division has been introduced in terms of a set of language-independent semantic primitives (semantic classes) assigned to all the nouns and verbs in the resource (Miller et al. 1993a). The verbs fall into 15 groups (Fellbaum 1998a: 70 – 71), such as verb.change (verbs describing change in terms of size, temperature, intensity, etc.), verb.cognition (verbs of mental activities or processes), verb.motion (verbs of change in the spatial position), verb.communication (verbs describing communication and information exchange), etc.²

Verb synsets are interrelated and form a hierarchical structure according to a troponymy relation (corresponding to hyponymy among nouns and representing a ‘manner’ relation); for example, in *talk – whisper*, the second member of the pair refers to a particular, semantically more specified, manner of performing the action referred to by the first verb (Fellbaum 1999: 94).

Communication verbs in WordNet are labelled with the semantic primitive verb.communication and cover primarily hyponyms (troponyms) of the synset {communicate, intercommunicate}, ‘transmit thoughts or feelings’.

WordNet is constructed according to principles governing human lexical memory in that it organises lexical information in terms of word meanings, rather than word forms, and uses an inheritance system reflecting a psycholinguistic judgement about the mental lexicon (Miller et al. 1993a: 14). The language-independent structure of word meanings allows the creation of wordnets for various languages linked to the Princeton WordNet through a set of interlingual indices, in essence representing a multilingual lexical semantic resource with comprehensive semantic description and a rich set of semantic relations.

WordNet provides extensive lexical coverage; the verbs represented in it are organised in 13,766 synsets (with additional 337 verb synsets specific for Bulgarian).

¹ <https://wordnet.princeton.edu/>

² The division of the nouns and verbs into WordNet lexicographic files (reflecting the semantic primitive distinction) along with short definitions of the primitives is available at: <https://wordnet.princeton.edu/documentation/lexnames5wn>.

However, it does not offer information on the realisation of the predicates and their arguments, and the syntactic information is limited to generic sentence frames illustrating simple sentences in which the verbs in the synset can be used (Miller et al. 1993a: 55, 80).

FrameNet³ (Fillmore 1982; Baker et al. 1998; Baker 2008) is a lexical semantic resource which couches lexical and conceptual knowledge in the apparatus of frame semantics. Semantic frames are conceptual structures describing types of objects, situations, or events along with their components (frame elements). Frame elements (FEs) may be core, peripheral or extra-thematic (Ruppenhofer et al. 2016: 23 – 24). In terms of the conceptual description, we deal primarily with core FEs, which instantiate conceptually necessary components of a frame, and which in their particular configuration make a frame unique and different from other frames.

FrameNet frames are organised into a hierarchical network by means of a number of hierarchical and non-hierarchical frame-to-frame relations (Ruppenhofer et al. 2016: 79 – 85). Here we list the hierarchical relations, which bear most relevance to the internal structure of thematic verb classes. These are: **Inheritance** – a relationship between a parent frame and a more specific (child) frame, such that the child frame elaborates the parent frame; **Uses** (also called ‘weak inheritance’) (Petrucci 2015: 33) – a relationship between two frames where the first one makes reference in a very general kind of way to the structure of a more abstract, schematic frame; **Perspective** – a relation indicating that a situation viewed as neutral may be specified by means of perspectivised frames that represent different possible points-of-view on the neutral state-of-affairs; **Subframe** – a relation between a complex frame referring to sequences of states and transitions, each of which can itself be separately described as a frame, and the frames denoting these states or transitions.

FrameNet also offers a set of annotated examples for lexical units evoking the corresponding frames (Ruppenhofer et al. 2016: 7 – 8). The annotation includes the verb, the frame elements and the syntactic component through which the frame element is realised. The annotation provides information both for the explicit and the implicit frame elements (non-overt but conceptually present frame elements retrievable from the immediate or the more general context, so-called null instantiations) (Ruppenhofer et al. 2016: 28 – 29; Petrucci 2019: 121). The set of examples supplies empirical evidence about the syntactic realisations of frame elements particularly valuable not only for linguistic generalisations about the target language (English) but as a point of departure for making observations cross-linguistically. However, FrameNet does not explicitly define the relevant selectional restrictions imposed by predicates on each frame element and its coverage is limited both in terms of the lexical units included in the frames (i.e. there are lexical units pertaining to a frame that are not listed in it) and in terms of the parts of the lexicon encompassed by the system of frames, i.e. there are many

³ <https://framenet.icsi.berkeley.edu/fndrupal/>

lexical units that cannot be described properly by the existing frames (Koeva 2020: 11 – 12).

Frames covering the class of communication verbs are related to the prototypical frame **Communication** and are presented in more detail in section 5.

2.2. Corpora

In order to explore the syntactic expression of the verbs and their participants, we study the usage examples available in two semantically annotated corpora – the English SemCor and the Bulgarian semantically annotated corpus, BulSemCor, both of which are annotated with WordNet senses. SemCor (current version 3.0) (Miller et al. 1993b, 1994; Landes et al. 1998) is compiled by the Princeton WordNet team and covers texts excerpted from the Brown Corpus. SemCor is supplied with part-of-speech and grammatical tagging and all open class words (both single words and multiword expressions, as well as named entities) are semantically annotated by assigning each word a unique WordNet sense (synset ID). The corpus is the largest manually annotated corpus of this kind and amounts to a total of 226,040 sense annotations.

BulSemCor (Koeva et al. 2006, 2011) is designed according to the general methodology of the original SemCor and criteria for ensuring an appropriate coverage of contemporary general lexis. The size of the corpus is close to 100,000 annotated units. The two corpora are not sufficient to provide enough evidence for many of the studied verbs so examples from other corpora have also been employed.

The Bulgarian-English Sentence- and Clause-Aligned Corpus (BulEnAC) (Koeva et al. 2012a) is a parallel corpus of aligned Bulgarian and English sentences and clauses with annotation of the syntactic relation between clauses. The corpus contains 366,865 tokens (176,397 tokens in Bulgarian and 190,468 tokens in English). The syntactic annotation of BulEnAC involves: a) sentence and clause splitting; b) annotation of the type of syntactic relation (coordinate or subordinate) between clauses. c) marking of the elements that introduce the clause: conjunctions, complementisers, and punctuation. BulEnAC is suitable for extracting parallel sentences illustrating the use of particular verbs evoking the frames under study. Further, it facilitates the identification of corresponding translation equivalents within aligned clauses.

The Bulgarian National Corpus is the largest corpus for Bulgarian: it consists of a monolingual (Bulgarian) part and 47 parallel corpora and amounts to 5.4 billion words. The Bulgarian part includes about 1.2 billion words of running text distributed in 240,000 text samples. The texts in the corpus reflect the state of the Bulgarian language predominantly in its written modality from the middle of the 20th century (1945) until the present day (Koeva et al. 2012b).

BulEnAC does not have semantic annotation, and the annotation in the Bulgarian National Corpus is supplied with all possible meanings of each verb, so the use of these resources requires disambiguation of the selected examples

by assigning the verb to a particular WordNet synset. Further, to enable the analysis of the semantic properties of frame elements, noun phrases also require disambiguation.

3. Related work

There are several other resources relevant to our study. VerbNet (Kipper-Schuler 2005; Kipper et al. 2008) provides substantial coverage of the English verb inventory and defines syntactic-semantic relations in an explicit way by means of predicate-argument structures. However, the existing mappings between WordNet synsets and VerbNet classes is limited.

VerbAtlas (Fabio et al. 2019) is a lexical-semantic resource representing the semantic description of the verb synsets in BabelNet. BabelNet is a very large, richly populated multilingual semantic network (covering more than 500 languages) which integrates lexicographic and encyclopaedic knowledge from WordNet and Wikipedia (Navigli, Ponzetto 2010). Each verb synset in VerbAtlas is assigned a frame corresponding to its prototypical predicate-argument structure. Obligatory components are described using 26 semantic roles and the semantic properties governing their compatibility (116 types).

Predicate Matrix (de Lacalle et al. 2014) is a lexical resource resulting from the integration of several sources of predicate information: FrameNet, VerbNet, PropBank and WordNet, that have been previously aligned in SemLink (Palmer 2009). Predicate Matrix is compiled using advanced graph-based algorithms to extend the mapping coverage between resources.

FrameNets have been developed for many languages, including Bulgarian. The original concept of the Bulgarian FrameNet was laid out by Koeva and Dekova (Koeva, Dekova 2008) and Koeva (2010) and further elaborated and implemented in later work (Koeva, Doychev 2022), resulting in the design of BulFrameNet – a web-based system for the comprehensive description of the semantic and the syntactic properties of verbs determining their syntactic realisation in text.

Combining the semantic description of verbs from different resources to enhance their representation has been proposed by Uresova et al. (Uresova et al. 2020a, 2020b). The result is a multilingual dictionary encoding a comprehensive description of the semantic classes of verbs and the semantic roles and syntactic properties of their arguments⁴. The project also aims at creating an ontology of events, processes and states, and for this purpose each dictionary entry is linked to its correspondences in FrameNet, WordNet, VerbNet, Ontonotes and PropBank, as well as the Valence Dictionary of Czech Verbs (Lopatkova et al. 2016), which represents the predicate-argument structure of each verb, its semantic class and the syntactic transformations (diatheses) in which it participates.

⁴ <https://ufal.mff.cuni.cz/synsemclass>

It has also been acknowledged that combining WordNet (and lexical semantic features) with resources such as FrameNet results in a more comprehensive semantic and syntactic representation of the lexical entries, thus expanding the possible applications of the resources for the purposes of syntactic and semantic parsing (Baker, Fellbaum 2009; Schneider et al. 2012; Das et al. 2014). A discussion on the strengths and shortcomings of the different kinds of lexical semantic resources has been offered by Shi and Mihalcea (Shi, Mihalcea 2005).

The possible alignment and mutual enrichment of the two resources have been the focus of research in the past decades (Baker, Fellbaum 2009; Tonelli, Pighin 2009; Palmer 2009; Laparra, Rigau 2010; Palmer et al. 2014; Leseva, Stoyanova 2020, among others).

One of the challenges in mapping resources developed according to different methodologies is the coverage of the alignment between the units represented in them. Our mapping uses as a point of departure three previously developed sources of lexical mappings: direct mappings provided within FrameNet (Baker, Fellbaum 2009), eXtendedWordFrameNet (Laparra, Rigau 2010), and MapNet (Tonelli, Pighin 2009), complemented with additional indirect mapping through VerbNet (Palmer 2009; Palmer et al. 2014). This resulted in 4,306 unique WordNet synsets to FrameNet frame mappings, achieving coverage of 30.5% out of all verb synsets (Leseva, Stoyanova 2020: 110).

Methods have been proposed to increase the coverage by discovering suitable literals based on semantic relations with literals already described in semantic frames (Burchardt et al. 2005) or by applying graph-based algorithms to identify relevant senses of verbs evoking certain semantic frames (de Lacalle et al. 2014). The main procedure we apply to improve and extend mapping coverage is based on exploration of the structural properties of WordNet and FrameNet. Verbs in a WordNet synset generally exhibit the same or very similar meaning, which implies that they are associated with the same semantic frame. Moreover, both resources are hierarchically organised based on the notion of inheritance from a more general to a more specific synset or frame. The alignment between the resources reflects the notion of inheritance – in general, more specific concepts should be associated with the frame of their hypernym(s) or with more specific frames elaborating on (and possibly inheriting from) this frame, although in practice this is not borne out consistently. The procedure we apply involves: (1) manual assignment of semantic frames to root verb synsets; (2) automatic assignment of the hypernym's frame onto hyponyms which were not previously mapped; and (3) validation and improvement of assignments with respect to precise and accurate representation of the situation. Using these procedures we have gradually increased the coverage of the mapping – achieving 94% coverage of the automatic mapping (Leseva, Stoyanova 2020: 115 – 116). It should be noted, however, that due to different reasons, such as specifics in the structure of WordNet, or lack of appropriate frames in FrameNet (where part of the lexis has not yet been described by frames) the automatic assignment needs expert verification. We thus have performed

manual validation, so far covering almost 50% of the mapping. The relatively language-independent description and the organisation principles underlying the design of both WordNet and FrameNet allows for the transfer of information both between resources and across languages.

The rich frame-to-frame relations employed in the structure of FrameNet are also reflected in the greater granularity and specification of the frame elements (as compared with VerbNet, VerbAtlas and other resources). The granularity is handled, where necessary, by applying a shallow hierarchy derived from the hierarchical organisation of the frames and the inheritance relations defined between them (Litkowski 2014). For example, the taxonomy of frame elements **Communicator** > **Speaker** > **Interlocutor** is derived from the frame hierarchy **Communication** > **Statement** > **Chatting** built on the frame-to-frame relation of weak inheritance (FrameNet relation **Uses**) between the three frames. The **Communicator** is the most generally defined frame element describing the agentive participant involved in either a verbal or non-verbal act of communication; the **Speaker** is constrained to a participant in spoken communication; the **Interlocutor** is more specific and refers to cases where the **Speaker** is one of a group in which the participants alternate between the roles of a **Speaker** and an **Addressee**. By maintaining different levels of granularity, FrameNet frames provide a more robust semantic description that is relatively resource- and theory-independent. In particular, this enables us to identify and refer to the frame elements' counterparts in other resources, as well as to adopt a more specific description or to resort to a more general one, depending on the particular task.

4. Language-independent semantic features in conceptual resources

The procedures for mapping FrameNet frames to WordNet synsets are based on the inheritance of semantic features in hypernym trees (Leseva et al. 2019: 281 – 282; Leseva, Stoyanova 2020: 110 – 111). In particular, we manually assign frames on root verb synsets and then rely on the assumption that the hyponyms either inherit the semantic frame of their hypernym directly or further elaborate on the frame of the hypernym.

Figure 1 illustrates a hypernym—hyponym pair of synsets, with the appropriate FrameNet frames assigned to them, which are themselves related by means of an inheritance relation (**Questioning** being an elaboration of the parent frame **Communication**).

Further, FrameNet frame elements are supplied with general semantic types (e.g., Sentient, Physical object, Time, etc.) defining the general noun semantic classes that may take the positions of the respective frame elements in the semantic frame (Ruppenhofer et al. 2016: 85 – 93). The semantic classes as part of the conceptual description are largely language-independent and can be transferred cross-linguistically.

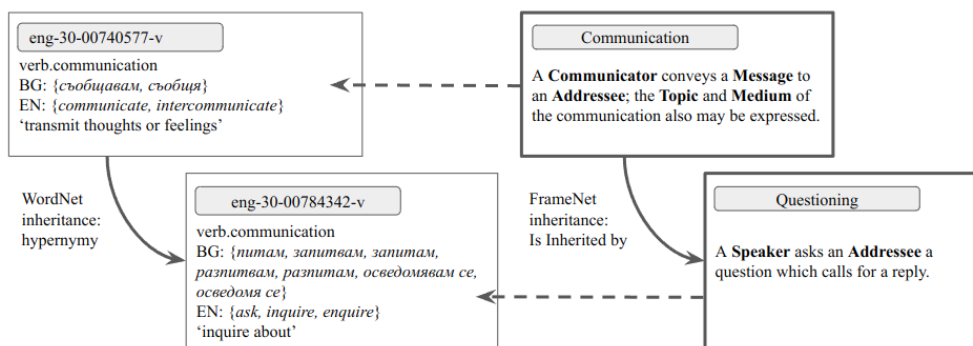


Figure 1. Frames inheritance (**Communication** → **Questioning**) reflected in synset hypernym / hyponym relations (*communicate* → *ask*)

The semantic classes can be modelled as a taxonomy using the WordNet structure (Koeva 2010: 48 – 51); they can be defined as (a combination of) WordNet substructures, i.e. hypernym-hyponym trees.

FrameNet provides a number of possible valence patterns for each lexical unit evoking a frame. While the general configuration of core frame elements could be proven to be to a great degree language-independent, the syntactic properties of the verbs and their realisation (as represented by the set of valence patterns) are more language-specific.

This work is grounded in the following assumptions: (a) cross-resource correspondence of facets of the semantic description stemming from similar semantic generalisations; (b) cross-lingual correspondence of semantic description; (c) relative cross-lingual syntactic correspondences of frame element realisations.

The first aspect is carried out through the linking of WordNet synsets and FrameNet frames whereby the synsets are assigned a schematic representation of the situation, the elements involved and their relationship to each other and the situation as described by the frame.

The second aspect is realised through the transfer of the semantic frame and valence patterns assigned to English verbs onto their counterparts in Bulgarian (using the linking between the Princeton WordNet and the Bulgarian WordNet).

The third aspect concerns the information about the lexical units evoking the FrameNet frames as reflected in the annotated examples in the FrameNet corpus. These reflect Fillmore's theory (Fillmore 1982: 376) on distinguishing 'case frames' as the structures in actual individual sentences in which the verbs could appear from 'case frame features' as representations of the class of 'case frames' into which particular verbs could be inserted, where 'cases' can be obligatory or optional and associated with some selectional dependencies. Each example for the realisation of a lexical unit in the FrameNet corpus of examples is supplied with annotation of the syntactic components expressing the relevant frame elements. In such a way

additional information is obtained about: (i) the combinatorial properties observed for a lexical unit through the so-called valence patterns, i.e. the particular combinations of both core and non-core frame elements that actually occur in the examples; (ii) the syntax-semantic interface, or the regularities in the syntactic projection of frame elements into the syntactic positions in clauses/sentences; (iii) the syntactic groups whereby each frame element is realised, along with their morpho-syntactic properties, lexical items that introduce them (prepositions, complementisers), etc.

The semantic frames are relatively invariable cross-linguistically in terms of the semantic information as they are grounded in human cognition and the conceptualisation of situations. To some extent generalised valence patterns are also valid across many languages. Observed variations cross-lingually are even more interesting as they point to potential language-specific properties of individual lexical units that may capture important cross-lingual contrasts. While being much more language-specific, the level of the syntax-semantics interface may be used as a point of departure in observing the syntactic realisation within and across languages, especially with a view to cross-lingual comparison and insights into the causes for variation. In FrameNet the projection of frame elements into syntactic positions is implemented in a straightforward manner by associating each frame element with a syntactic category that is further specified for its grammatical function – e.g. subject (NP.Ext) and object (NP.Obj). This declarative linking enables the direct observation of the syntactic properties and behaviour of lexical units.

While there certainly are mismatches in the syntactic categories whereby frame elements are expressed across languages, differences are found particularly at the level of grammatical features, categories and constructions. For instance, Bulgarian lacks *-ing* and infinitive clauses, so propositional complements will be realised as finite clauses; Bulgarian has impersonal verbs so there are impersonal sentences, whose English counterparts would involve a dummy subject, etc. In addition, languages may also differ in terms of the overtness of syntactic information, i.e. the possibility to leave an obligatory element non-explicit (null instantiations retrievable from the context or the grammatical construction); the language-specific diatheses, constructions, word order, morphosyntactic features, etc. The specific inventories of linguistic means serving to introduce certain frame elements such as prepositions, conjunctions, *wh*-words, etc. is to a great degree language-specific.

The semantic and syntactic annotation undertaken within the FrameNet project provides valuable observational data, whose validity we test for the Bulgarian counterparts of the verbs in the selected frames as represented in the dataset of annotated Bulgarian examples. In particular, we check the applicability of the valence patterns attested for the English verbs included as lexical units in FrameNet to their Bulgarian counterparts. This is possible due to the fact that part of the English synsets (and by virtue of the interlingual correspondence between equivalent senses, the Bulgarian synsets as well) have been mapped to FrameNet frames (Leseva, Stoyanova 2020). The analysis of valence patterns may be extended to members of Princeton WordNet synsets and their Bulgarian counterparts which do not have a correspondence in FrameNet but have been

assigned a suitable frame automatically within some of the efforts at aligning FrameNet and WordNet described in Section 3. This is particularly applicable to the analysis of examples extracted from SemCor, BulSemCor or other annotated resources, where other verbs evoking a given frame but not included in FrameNet may be found.

For instance, the synset {announce, denote}, ‘make known; make an announcement’ is assigned the semantic frame **Statement** by virtue of the fact that the synset member *announce* has a counterpart with a matching sense among the lexical units in FrameNet evoking the frame **Statement**. By extension, initially the same frame has been assigned to the synset {blare out, blat out}, ‘announce loudly’, neither member of which has a matching lexical units in FrameNet⁵, through an automatic procedure making use of the fact that it is a hyponym of {announce, denote} and shares its central meaning component. Our assumption⁶ is that since the meaning of the hyponym is a concretisation of the meaning of the hypernym, then the frame evoked by the hyponym will either be the same as that of the hypernym, or be a more concrete frame related to the frame of the hypernym. From this assumption, we proceed to analyse whether the sense of this synset is indeed appropriately described by this semantic frame and whether the syntactic realisation of its members follows the valence patterns typical for the frame. Once the validity of assignment of the semantic frame is established, using the correspondence between the equivalent synsets in the Princeton WordNet and BulNet, the assignment of the **Statement** frame to {blare out, blat out} is straightforwardly transferable to the respective Bulgarian synset and to the instantiations of its members in the Bulgarian annotated dataset.

In examining the semantic and syntactic properties of individual verbs, we take into account both the most frequent valence patterns for the relevant frame obtained by aggregating the valence patterns for all the verbs that evoke it (Table 1) and the valence patterns typical of the individual verb, or if it does not have a counterpart in FrameNet, its closest match.

⁵ The verb *blare* is found as a lexical unit evoking the frames **Make noise** (as in FN: *Some-where behind her a horn blared*) and **Cause to make noise** (FN: *A nearby car blared its horn loudly*), but these frames involve the production of noises and sounds without a communicative aspect.

⁶ The idea has been already introduced by Koeva (2020: 15 – 16): “Inheritance is important in the way that all noun synsets that are hyponyms of a synset representing a particular semantic class should inherit the properties of this class, and also all verb synsets that are hyponyms of a synset associated with a particular conceptual frame should inherit the properties of this frame ... As for the conceptual frames (if they are correctly defined) within a fine-grained WordNet structure of inheritance relations we can expect that the daughter verb synsets will inherit the conceptual frame assigned on the top of the tree and deviations are expected in two directions: a reduction of a core frame element and a reduction of the members of the set of nouns eligible to express a particular frame element”.

| Pattern | # | verbs |
|---|-----|---|
| [NP.Ext] _{Speaker} [Clause] _{Message} | 281 | <i>declare, remark, announce</i> |
| [NP.Ext] _{Speaker} [NP.Obj] _{Message} | 191 | <i>explain, note, declare, announce</i> |
| [NP.Ext] _{Speaker} [Quote] _{Message} | 143 | <i>explain, insist, propose, announce</i> |
| [NP.Ext] _{Speaker} [PP] _{Topic} | 83 | <i>explain, gloat, preach, report</i> |
| [NP.Ext] _{Medium} [Clause] _{Message} | 39 | <i>note, declare, allege, announce</i> |

Table 1. A sample of the most frequent valence patterns in the frame **Statement**

The comparison between the aggregated valence patterns for the semantic frame and the individual valence patterns for a given verb shows the similarities and differences between the observations for the verbs in the frame in general and the observations for the particular verb. In Section 5 we present a case study on communication verbs and illustrate their general semantic and syntactic properties and behaviour as represented in the FrameNet corpus and the Bulgarian annotated dataset. In Section 6 we zoom in on several high frequency Bulgarian verbs evoking the frames **Statement** and **Telling**. We analyse their semantic and syntactic features, pointing out certain specificities, in comparison with each other and against the characteristics of their English counterparts retrieved from the data in the FrameNet corpus.

5. A case study: verbs of communication

Below we discuss verbs of communication with the objectives to: (a) validate the conceptual description available in the lexical-semantic resources against corpus data for Bulgarian; and (b) demonstrate the language-independent principles and the cross-lingual validity of the semantic frames **Communication**, **Statement** and **Telling**, the attested valence patterns and (partly) the syntactic description of frame elements for the pair English-Bulgarian. Speech act verbs have been discussed by many authors with a particular focus on their semantic and syntactic properties and classification (Wierzbicka 1987; Levin 1993: 202 – 211; Levin et al. 1997; Urban, Ruppenhofer 2001; Boas 2010, among others).

Our observations are performed on a dataset of 4,525 annotated illustration examples containing verbs of communication representing 863 different valence patterns in English extracted from FrameNet, and additional 890 manually

annotated sentences representing 136 different valence patterns in Bulgarian. The annotation of the Bulgarian sentences follows the FrameNet annotation guidelines (Ruppenhofer et al. 2016: 19 – 52) and consists in labelling the sentence components with the frame elements they realise. We should also note that the data presented reflect the distribution in our datasets. Some valence patterns that are missing from the data for Bulgarian are actually represented in the language but are either less frequent or the selection of examples may have been skewed towards certain patterns.

5.1. Organisation of FrameNet frames within the class of verbs of communication

The prototypical frame characterising the studied domain is the frame **Communication**. It describes the general scenario of people or entities interacting to the end of imparting or exchanging information in terms of the participants and circumstances involved and the relations between them. This scenario is then elaborated in various ways in more specific frames. The frames in this class form a shallow hierarchical structure branching from **Communication** mainly through the frame-to-frame relations of inheritance (FrameNet relation **Inherits from**) and weak inheritance (FrameNet relation **Uses**). The structure is shown in Figure 2.

For each such frame discussed below (**Communication**, **Statement** and **Telling**), we present: (i) its semantics in terms of the semantic frame definition, constellation of core frame elements that represent the main participants in the situation, and the relations among them as described in FrameNet, together with (ii) the syntactic expression of the frame elements in English (according to the annotation in the FrameNet corpus) and Bulgarian (as observed in the annotated dataset), and (iii) the specifics of their realisation in Bulgarian as compared to English.

5.2. The prototypical frame of the communication domain: Communication

5.2.1. General description

Communication is the prototypical frame that represents the basic conceptual structure of the activity of communication. It describes a directed act of communication where a **Communicator** produces a **Message** (on a **Topic**) and directs it towards an **Addressee**⁷. It includes various forms of communication,

⁷ The frame elements are defined as follows: **Communicator**: “The sentient entity that uses language in the written or spoken modality to convey a **Message** to the **Addressee**.”; **Medium**: “The physical or abstract setting in which the **Message** is conveyed.”; **Message**: “**Message** is a proposition or set of propositions that the **Communicator** wants the **Addressee** to believe or take for granted.”; **Topic**: “The **Topic** is the subject matter to which the **Message** pertains. It is normally expressed as a PP Complement headed by “about”, but in some cases it can appear as a direct object.”; **Addressee** (non-core): “The **Addressee**

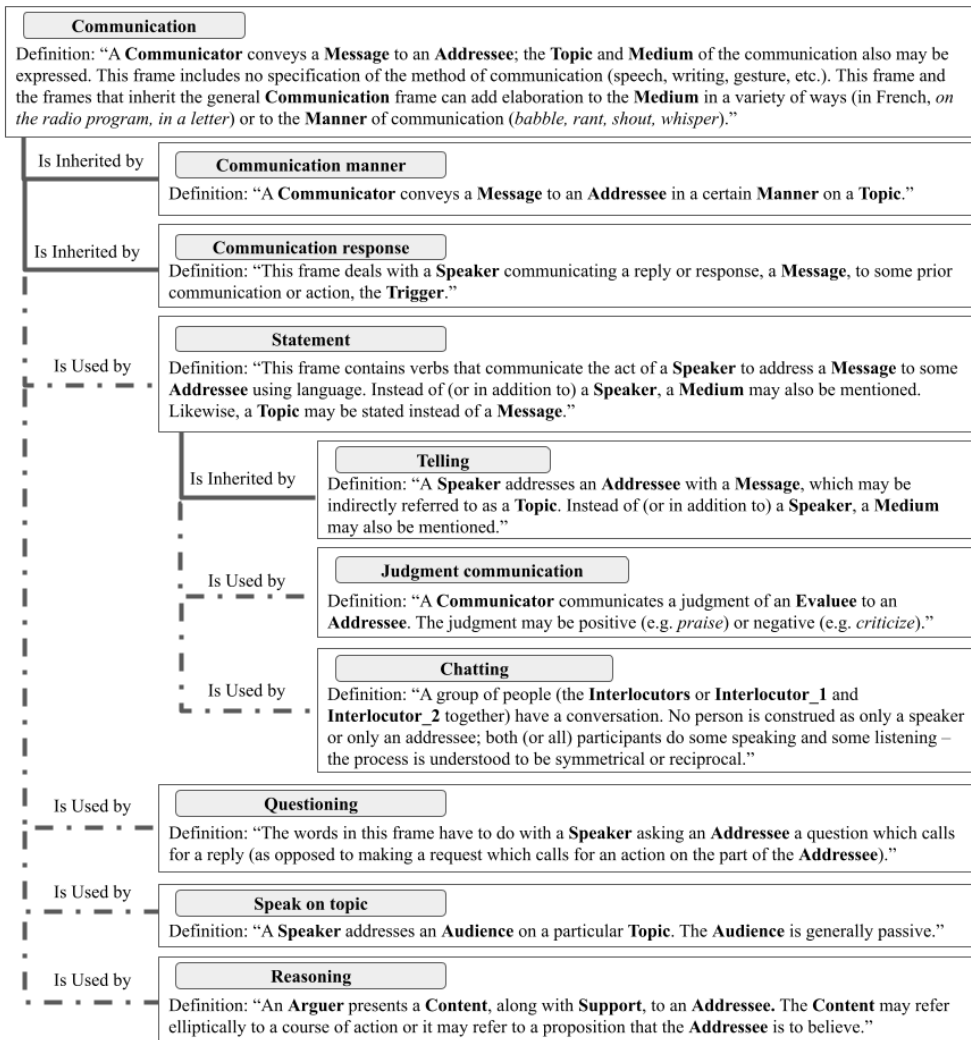


Figure 2. Hierarchical structure of the semantic frames describing communication verbs

both verbal and non-verbal. This basic structure is further elaborated (narrowed down, profiled or otherwise specialised) in the frames that inherit it. The FrameNet frame **Communication** is evoked by a limited number of verbs – *communicate, convey, indicate, share*. Although pertaining to the prototypical frame, these verbs

receives a **Message** from the **Communicator**". For the relevant definitions in the remaining frames discussed herein, see the respective representation in FrameNet: https://framenet.icsi.berkeley.edu/framenet_search.

are not the most frequent ones associated with the activity of communicating, which are in fact described in more elaborate semantic frames inheriting from **Communication**, such as **Statement**, **Telling**, **Communication noise**, etc.

The prototypical frame presents a generalised communication scenario elaborated by more concrete semantic frames in the following ways:

- (i) by limiting the scope of the method of communication. For instance, while the frame **Communication** covers the general scenario of either verbal or non-verbal communication, some frames are constrained in this respect, e.g. **Statement**, evoked by verbs such as *say*, *state*, *declare*, involves speech;
- (ii) by specifying the instrument (**Means**), i.e. the physical entity, channel, medium, form or technology required or involved to carry out the communication, as in the frame **Communication means** (e.g., *fax*, *phone*, *radio*);
- (iii) by specifying the manner of verbal communication according to various criteria such as loudness (e.g., *shout*, *whisper*); volubility and/or mood (e.g., *babble*, *rant*), distinctness (e.g., *slur*, *stutter*, *mutter*), among many others (**Communication manner**);
- (iv) by elaborating or specialising on the meaning of the **Communication** frame in such a way that the more concrete frames inherit only part of its frame elements, do not inherit them in a straightforward manner or introduce additional frame elements. For instance, the frame **Judgment communication**, which Uses **Statement**, does not conceptualise the **Message**. The lexical units evoking this frame, such as the verbs *criticise*, *praise*, *belittle*, express positive or negative judgement, which is incorporated in the lexical unit and may be viewed as a proxy of the message. Besides, the frame introduces the frame elements **Evaluee**, i.e. the object being judged, **Expressor** (the body part expressing the judgement) and **Reason** (the state-of-affairs that describes the justification or reason for the judgement);
- (v) by narrowing down the semantic class of the frame elements. For instance, in some frames the **Communicator** is specialised as a **Speaker** who delivers a message verbally (e.g., the frame **Statement**), or as an **Interlocutor**, who switches between the roles of a speaker and an addressee (e.g., the frame **Chatting**).

The incorporation of frame elements (Jackendoff 1990: 164 – 165) is a specific type of frame specialisation whereby a certain frame element is integrated in the meaning of a verb as a result of which this frame element is usually left unexpressed (Ruppenhofer et al. 2016: 30). For example, the frame **Communication means** describes situations that are characterised by concrete means with the aid of which communication takes place; the various **Means** are incorporated in the meaning of the respective verbs, e.g. *fax*, *telephone*, *email*.

5.2.2. Syntactic realisation of the frame elements in the **Communication frame**

Below we describe the syntactic realisation of the most essential frame elements in the **Communication** frame as both their definition and their syntactic characteristics and behaviours determine to a great degree the specificities of the frame elements in the more concrete frames. Koeva (Koeva 2020) gives a general overview of the syntactic realisation of verb frame elements and the representation of syntactic information in various semantic resources. Koeva (Koeva 2019: 60 – 61) presents the complements that transfer of information verbs take in Bulgarian, among other verb classes. These are relevant for the analysis of the syntactic realisation of frame elements in the frames evoked by communication verbs.

As the annotated FrameNet examples show, the **Communicator** is expressed as the subject of the respective sentence or clause. According to the annotation conventions adopted in the resource (Ruppenhofer et al. 2016: 71 – 72), the subject is marked with the label NP.Ext (standing for external argument). The frame element may be syntactically unexpressed. In such cases, it is annotated as a definite null instantiation (DNI) if its referent is retrievable from the previous context, or as a constructional null instantiation (CNI), when it remains implicit by virtue of the grammatical construction, e.g. in passive or infinitive clauses.

The analysis of the FrameNet valence patterns shows that the **Message** is realised as a direct object (NP.Obj) (1.c.), as a complement clause (Clause) (1.a.) or as a quote (Quote) (1.e.). Quotes represent the content of the **Message** as directly stated by the **Communicator** in their own words, while clauses denote it as being retold by someone (as reported speech). A **Message** realised as a direct object constitutes a nominalisation which rephrases its content in a more concise way or as a generalised idea. In a considerable portion of the cases (around a third) the **Message** is annotated as an indefinite null instantiation (INI). This means that the verb is used intransitively, the **Message** remains syntactically unexpressed and receives a certain typical interpretation without a specific discourse referent (Ruppenhofer et al. 2016: 19 – 21).

The FrameNet examples show that the **Topic** is rarely expressed, with a limited number of annotated examples for all of the studied frames. It is usually realised as a prepositional phrase most often headed by ‘about’, but this may vary from verb to verb (1.b., 1.d.). An alternative way of realising the **Topic** is as a modifier of a noun expressing the **Message**; such cases corroborate syntactically its semantic dependence on the **Message** communicated. In the absence of an overt **Message**, the **Topic** may be expressed as an independent phrase (1.d.); this is one of the typical valence patterns of its realisation as attested in the more specific communication frames. Rarely, such a phrase appears in the presence of a **Message** (1.e.).

The **Medium** is expressed either as a prepositional phrase, or as the subject in the case of a non-overt **Communicator**.

The **Addressee** is either realised as a prepositional phrase or is left unexpressed, although its presence is always required conceptually as any act of communication

is directed to someone. Predominantly, the non-overt **Addressee** frame elements are marked as indefinite null instantiations (INI).

- 1.a. [*Iranian officials*]_{Communicator} **INDICATE** [*that Iran would honor its safeguards agreement with the IAEA*]_{Message} []_{Addressee-INI}.
- 1.b. [*They*]_{Communicator} *can easily* **COMMUNICATE** []_{Message-INI} [*with one another*]_{Addressee}.
- 1.c. [*The letter*]_{Communicator} **COMMUNICATED** [*nothing*]_{Message} [*of her pleasure*]_{Topic}.
- 1.d. [*I*]_{Communicator} **COMMUNICATED** [*with the Minister*]_{Addressee} [*on that issue*]_{Topic}.
- 1.e. [*“That was an incredible experience!”*]_{Message} [*he*]_{Communicator} **SHARED** []_{Addressee} [*about the trip*]_{Topic}.

The following clarifications should be made. FrameNet accounts for the conceptual interdependence observed among some semantic arguments of verbs, specifically the fact that “the presence of any member of the set is sufficient to satisfy a semantic valence of the predicator” by grouping the relevant frame elements in the so-called coreness sets, or Core Sets (Ruppenhofer et al. 2016: 25 – 26). The correlations in the syntactic expression of different frame elements outlined above follow from such semantic dependencies. In particular, two Core Sets are specified in the **Communication** frame: {**Communicator, Medium**} and {**Message, Topic**}. The former one {**Communicator, Medium**} captures the generalisation that, while the **Medium** is the “setting” used by the **Communicator** to deliver the **Message**, some types of **Medium** may be construed as the sources or expressors of the **Message** independently in the absence of an expressed **Communicator**, who is nonetheless conceptually present. This enables the realisation of either or both elements in the set. The latter Core Set accounts for the intrinsic dependence that holds between the **Topic** and the **Message**, stemming from the fact that the **Topic** is the subject matter of the **Message** and therefore it is always predicated on the existence of a **Message**. This dependence presupposes the variants in the expression of the two frame elements, as discussed in Sections 5 and 6. Core Sets are also inherited among frames (although not always in a straightforward manner, see Ruppenhofer et al. 2016: 27), reflecting the preservation of the semantic dependencies among frame elements. In the frames discussed below this observation is fully confirmed.

In Bulgarian the core frame elements are expressed in similar configurations of frame elements realised by means of mostly the same syntactic categories as in English (Table 2) as illustrated in Example 2. In particular, the **Communicator** is most frequently the subject; the **Message** is realised as an NP object (2.a. – 2.c.) or more rarely (although varying from verb to verb) as a complement clause (2.d.) or a quote (2.e.). When overt, the **Addressee** is expressed as a prepositional phrase (2.b.). Example (2.a.) shows the **Topic** realised as a prepositional phrase that modifies the **Message** head noun. However, as noted earlier, it may also be

expressed as an independent phrase even in the presence (but most often in the absence) of the **Message** (2.e., which is a translation of 1.e.).

2.a. [Te]_{Communicator} СЪОБЩАВАТ [съответната информация]_{Message} [за дейността си]_{Topic}.

They COMMUNICATE relevant-DEF information about activity-DEF REFL.POSS.

‘They communicate relevant information about their activity.’

2.b. [Te]_{Communicator} СЪОБЩАВАТ [на Комисията]_{Addressee} [текста на разпоредбите]_{Message}.

They COMMUNICATE to Commission-DEF text-DEF of measures-DEF.

‘They communicate to the Commission the text of the measures.’

2.c. [Органите]_{Communicator} СЪОБЩАВАТ [цялата информация]_{Message} []_{Addressee}.

Authorities-DEF COMMUNICATE all-DEF information.

‘The authorities communicate all essential information.’

2.d. [Страните]_{Communicator} ПОСОЧВАТ, [че поверителната информация не може да бъде резюмирана]_{Message} []_{Addressee}.

Parties-DEF INDICATE that confidential-DEF information cannot be summarised.

‘The parties indicate that the information cannot be summarised.’

2.e. [– Невероятно преживяване!]_{Message} – СПОДЕЛИ [той]_{Communicator} []_{Addressee} [за пътешествието си]_{Topic}.

– Quite an experience! – SHARED he about trip-DEF REFL.POSS.

‘‘That was quite an experience!,’’ he shared about the trip.’

The summary of the most frequent valence patterns attested for **Communication** verbs in the FrameNet corpus and in the Bulgarian annotated dataset (Table 2) shows the distribution of these patterns across verbs in the two languages with the corresponding number of examples. The corpus occurrences in both English and Bulgarian fall into several valence patterns that involve basically the **Communicator** and the **Message**, while the **Addressee** is more frequently non-overt. Valence patterns involving the **Medium** and the **Topic** are quite rare, possibly due to the relatively small dataset.

| Valence patterns | # | EN verbs | # | BG verbs |
|---|----|--------------------------------------|----|--|
| [NP.Ext] _{Communicator} [NP.Obj] _{Message} [PP] _{Addressee} | 11 | <i>communicate,</i> <i>signal</i> | 50 | <i>предавам / предам,</i> <i>споделям / споделя,</i> <i>съобщавам /</i> <i>съобщя</i> |
| [NP.Ext] _{Communicator} [PP] _{Addressee} [] Message-INI | 7 | <i>communicate</i> | | |

| | | | | |
|---|---|----------------------------------|---|--|
| [NP.Ext] _{Communicator} [] _{Addressee-INI} [NP.Obj] _{Message} | 5 | <i>communicate</i> | 9 | <i>споделям / споделя, предавам / предам</i> |
| [NP.Ext] _{Communicator} [] _{Addressee-INI} [] _{Message-INI} | 4 | <i>communicate</i> | | |
| [NP.Ext] _{Communicator} [Clause] _{Message} [] _{Addressee-INI} | 4 | <i>indicate, say, signal</i> | 2 | <i>съобщавам / съобща</i> |
| [NP.Ext] _{Message} [PP] _{Addressee} [] Communicator-CNI | 3 | <i>communicate</i> | | |
| [NP.Ext] _{Communicator} [] _{Addressee-DNI} [NP.Obj] _{Message} | 3 | <i>communicate, indicate</i> | 9 | <i>споделям / споделя, предавам / предам</i> |
| [NP.Ext] _{Communicator} [PP] _{Addressee} [NP.Obj] _{Message} [PP] _{Topic} | 2 | <i>communicate</i> | 1 | <i>съобщавам / съобща</i> |
| [NP.Ext] _{Medium} [] _{Addressee-INI} [Clause] _{Message} | 2 | <i>indicate</i> | | |

Table 2. FrameNet valence patterns of **Communication** verbs, with their frequency in the FrameNet corpus and the verbs they appear with, compared with the Bulgarian data

Table 3 summarises the generalised semantic classes of the nouns representing the frame elements with a particular syntactic realisation. Frame elements also form a shallow hierarchical structure based on the inheritance relations between the corresponding frames (as illustrated for the elements Communicator > Speaker > Interlocutor in Section 3), where the subordinate elements possess additional or more specific semantic properties.

| Frame | Frame element | Syntactic realisation | Most typical semantic class |
|-----------------------------------|---------------|-----------------------|-----------------------------|
| Communication | Communicator | NP | {person} |
| Statement, Telling | Speaker | NP | {person} |
| Telling | Addressee | NP | {person} |
| Communication, Statement | Addressee | PP | <i>to</i> + {person} |
| Communication, Statement, Telling | Message | NP | {message} |
| | | Clause | N/A |
| | | Quote | N/A |
| Communication, Statement, Telling | Topic | NP | {entity} |
| | | PP | <i>about</i> + {entity} |

Table 3. A summary of the semantic classes and the syntactic specificities of the main frame elements in the frames of the communication domain

5.3. Statement

5.3.1. General description

The semantic frame **Statement** inherits from the prototypical frame **Communication** via the weak-inheritance relation **Uses** and specifies communication involving the expression of a verbal (written or oral) message. In terms of the number of verbs it evokes, including many general lexis verbs, such as *say, state, declare, speak, report, note*, etc., it may be considered the most representative frame of verbal communication.

The core frame elements in terms of which the frame is described include: **Speaker, Message, Medium** and **Topic**. While conceptually implied in the act of communication, the **Addressee** is not specified as a core element of the frame. This reflects the fact that the semantic core of the lexical units evoking the frame **Statement** is considered to be the process or event of the **Message** being conveyed by the **Speaker**. This may be observed in the valence patterns common for the frame (Table 4), where the various combinations of the **Speaker** and the **Message** account for most of the frequent valence patterns (the **Message** can be substituted by its **Topic**). The (intended) receiver of the **Message**, i.e. the **Addressee**, is semantically backgrounded and thus considered optional and often left syntactically non-overt. When it is expressed, it is usually together with the **Message** (Table 4). Even when non-overt, we take the **Addressee** into account in the process of annotation.

5.3.2. Syntactic realisation of the frame elements in the Statement frame

The fact that **Statement** represents an elaboration of **Communication** involving language and the faculty of speech is reflected in FrameNet through the reconsideration of the frame element **Communicator** of the parent frame as the more specific **Speaker**, which denotes the person who produces the message. Likewise, the **Speaker** is realised as the external NP.

The **Medium** and the **Speaker** share the same relation of interdependence as the **Medium** and the **Communicator** in the **Communication** frame, forming a CoreSet, meaning that the expression of only one of them is sufficient to realise the semantic valence of the verb. This results in the possibility for the **Medium**, which in the presence of an overt **Speaker** is expressed as a prepositional complement, to occupy the subject position if the **Speaker** has an implied, non-specified reading and is left unexpressed (3.f.).

According to the data, the **Message** is most often expressed either as a subordinate clause, an NP object, or a direct quote rendering the content being conveyed (Example 3.a. – 3.c.). The possible realisations vary across verbs: some of them have a stronger tendency to take a complement subordinate clause (e.g., *claim, suggest, note*), while others favour an NP object (e.g., *profess, reiterate, relate*) or a quote (e.g., *exclaim*); in some cases the three realisations are relatively equally represented (e.g., *caution*).

The **Topic** is usually expressed as a prepositional phrase headed by various prepositions depending on the particular verb (e.g. *speak about him, speak of him, preach of heaven, comment on the protests, comment upon the economic conditions*). The nature of the relation existing between the **Message** and the **Topic**, as explained in the subsection on the **Communication** frame, is one of interdependence, which allows only one of the two frame elements to be present for an utterance to be semantically and grammatically felicitous. Indeed, in most examples in the dataset either the one or the other is expressed overtly, although the two may also co-occur. In the latter case, the **Topic** is usually syntactically dependent on the **Message** (3.b.). Though rarely, as attested for some verbs, it can even be projected as a direct object (3.d.). This is possible when the **Topic** stands for the **Message** in its absence.

When expressed, the **Addressee** is realised as a prepositional phrase usually headed by the preposition *to* (3.d.) or as an indirect object (3.e.); as shown in Table 4, the valence patterns involving this frame element are quite infrequent.

- 3.a. [*North Korea*]_{Speaker} CLAIMED [*it had no intention of producing nuclear weapons*]_{Message}.
- 3.b. [*He*]_{Speaker} SAID [*little*]_{Message} [*about the case*]_{Topic}.
- 3.c. [*He*]_{Speaker} ADDED: [*'Eldorado is a brave venture'*]_{Message}.
- 3.d. [*The doctor*]_{Speaker} EXPLAINED [*the injuries*]_{Topic} [*to the police*]_{Addressee}.
- 3.e. [*The agency*]_{Speaker} WROTE [*me*]_{Addressee} [*that you had moved*]_{Message}.
- 3.f. [*The letter*]_{Medium} ALLEGED [*serious breaches of the law*]_{Message}.

The syntactic realisation of the frame element configurations in Bulgarian closely resembles that in English. The **Speaker** is usually realised as the external NP and can be a person, a group or an organisation (4.a., 4.b.). In some cases the **Medium** can take the position of the external argument (4.c.).

The types of complements selected by communication verbs have been described by Koeva (Koeva 2019: 60 – 61). More specifically, she divides ‘transfer of information verbs’ into three groups according to their general meaning and complements: (i) verbs with a direct object addressee (these cover the verbs in the **Telling** frame); (ii) verbs with an indirect object addressee (introduced by the preposition *na* (*to*), which cover **Statement** verbs and possibly verbs evoking some other frame); (iii) verbs with an unexpressed addressee. In the first case, (i), the information transferred (i.e. the message) is expressed by a PP headed by the preposition *za* (*about*) and/or a complement *ue*-clause (*that*-clause) or an embedded *wh*-question; in the second, (ii), the message is rendered as an object NP alternating with a complement *ue*- or *wh*-clause or as a PP headed by *za* (*for, about*) and/or a complement *ue*- or *wh*-clause; in the third, (iii), the information transferred may be projected as a complement *ue*- or *wh*-clause alternating with an object NP, or as a PP headed by *za* (*for, about*) and a complement *ue*- or *wh*-clause. The valence pattern where the **Message** is realised as a *za*-PP and a complement clause is not represented in our dataset.

In this Section as well as in Sections 5.4, 6 and 7 we elaborate on these observations and look into the distribution and possibly at certain specifics of the different types of complements as compared with English.

The **Message** is likewise realised as a finite clause, an object NP or a direct quote (4.a., 4.b., 4.f.). The **Topic** (4.d.) and the **Addressee** (4.b.) have the same syntactic behaviour as expected: they are frequently non-overt, and when they are explicit, they are expressed as prepositional phrases. The **Topic**'s dependent or independent expression follows the same valence patterns as in English.

4.a. [*Панайотов*]_{Speaker} *ДОБАВИ*, [*че лидер на партията ще е Симеон*]_{Message}.

Panayotov ADDED that leader-DEF of party-DEF will be Simeon.

'Panayotov added that Simeon will be the leader of the party.'

4.b. [*Кредитните институции*]_{Speaker} *ДОКЛАДВАХА* [*на Алън Грийнспан*]_{Addressee} [*повишено ниво на покупки*]_{Message}.

Credit institutions REPORTED to Alan Greenspan increased level of purchases.

'Credit institutions reported to Alan Greenspan increase in purchases.'

4.c. [*Неофициалните статистики за 1999 г.*]_{Medium} *СОЧАТ* [*5000 посетители*]_{Message}.

Unofficial-DEF statistics for 1999 REPORT 5000 visitors.

'The unofficial statistics for 1999 report 5000 visitors.'

4.d. [*Тези лица*]_{Speaker} *ИЗКАЗВАТ* [*пред нас*]_{Addressee} [*неприятни истини*]_{Message} [*за смъртните ни врагове*]_{Topic}.

These people REPORT to us unpleasant truths about mortal-DEF our enemies.

'These people report to us unpleasant truths about our mortal enemies.'

4.e. [*В интервюто*]_{Medium} [*Симеон*]_{Speaker} *ОБЯВИ* [*промяна на политическата посока*]_{Message}.

In interview-DEF Simeon ANNOUNCED change of political-DEF direction.

'In the interview Simeon announced a change in the political direction.'

4.f. [*Тя го каза просто така*]_{Message} – *ДОБАВИ* [*Дженифър*]_{Speaker}.

She it said just so – ADDED Jennifer.

“She said it just like that,” added Jennifer.'

The various specific configurations of frame elements as expressed by verbs in the **Statement** frame are shown in Table 4.

| Valence patterns | # | EN verbs | # | BG verbs |
|---|-----|---|----|--|
| [NP.Ext] _{Speaker} [Clause] _{Message} | 281 | <i>explain, note, declare, maintain, remark, mention, conjecture, reiterate, assert, preach, claim, attest, state, caution, write, add, allege, exclaim, say, suggest, insist, propose, announce, confirm, acknowledge, proclaim, reaffirm, report, pronounce</i> | 67 | <i>казвам / кажа, добавям / добавя, коментирам, заявявам / заявя, обявявам / обявя, обяснявам / обясня, отбелязвам / отбележа, пиша, твърдя, посочвам / посоча, предлагам / предложа</i> |
| [NP.Ext] _{Speaker} [NP. Obj] _{Message} | 191 | <i>explain, note, declare, tell, conjecture, reiterate, assert, preach, claim, speak, talk, state, caution, write, add, al lege, exclaim, say, suggest, propose, announce, confirm, acknowledge, refute, proclaim, reaffirm, report</i> | 29 | <i>казвам / кажа, коментирам, обявявам / обявя, оповестявам / оповестя, повтарям / повторя, посочвам / посоча, предлагам / предложа, съобщавам / съобщя</i> |
| [NP.Ext] _{Speaker} [Quote] _{Message} | 143 | <i>explain, gloat, declare, remark, observe, mention, reiterate, hazard, assert, preach, speak, attest, state, caution, write, add, allege, exclaim, say, pout, suggest, insist, propose, announce, proclaim, reaffirm, report</i> | 48 | <i>казвам / кажа, добавям / добавя, коментирам, заявявам / заявя, обявявам / обявя, обяснявам / обясня, пиша, твърдя, отбелязвам / отбележа, повтарям / повторя, съобщавам / съобщя</i> |
| [NP.Ext] _{Speaker} [] Message-INI [PP] _{Topic} | 83 | <i>explain, gloat, preach, report, comment, remark, speak, talk, write</i> | | |
| [NP.Ext] _{Medium} [Clause] _{Message} | 39 | <i>note, declare, allege, say, suggest, propose, announce, confirm, acknowledge, proclaim, report, claim, state</i> | | |
| [NP.Ext] _{Speaker} [PP] Addressee [NP.Obj] _{Message} | 28 | <i>reiterate, declare, report, say, speak, state, suggest, propose, announce, mention</i> | 2 | <i>обяснявам / обясня</i> |
| [NP.Ext] _{Speaker} [PP] Addressee [Quote] _{Message} | | | 5 | <i>казвам / кажа, заявявам / заявя</i> |

| | | | | |
|---|----|--|---|---|
| [NP.Ext] _{Speaker} [PP] Message | 28 | <i>profess, declare, preach, say, speak, de scribe, insist, caution</i> | | |
| [NP.Ext] _{Speaker} [PP] Addressee [Clause] _{Message} | 25 | <i>add, explain, declare, allege, suggest, insist, propose, announce, mention, confirm, preach</i> | 9 | <i>заявявам / заявя, обяснявам / обясня, предлагам / предложи, съобщавам / съобщя</i> |

Table 4. FrameNet valence patterns of **Statement** verbs, with their frequency in the FrameNet corpus and the verbs they appear with, compared with the Bulgarian data

5.4. Telling

5.4.1. General description

The frame **Telling** is evoked by a small number of frequently occurring verbs such as *tell, advise, inform, notify*, etc. The frame inherits from **Statement** and its specialisation consists in describing the speech act as directed to a specific **Addressee**. This shift of focus from the production of the **Message** to the recipient who is addressed with the message, results in the promotion of the frame element **Addressee** to core status, and with most of the verbs (*inform, advise, confide, notify*) it is the one favoured for the direct object position.

The core frame elements are: **Speaker, Addressee, Message, Medium, Topic**. The frame elements generally have the same characteristics as the ones in the **Statement** frame from which they inherit their properties.

The most frequent valence patterns observed in the data are shown in Table 5 in a comparative perspective between English and Bulgarian. Some patterns, although rare (and thus not present in the annotated data for Bulgarian), are still possible (pattern [NP.Ext]_{Speaker} [NP.Obj]_{Addressee} [PP]_{Message?} 6.f.). More data are needed to ensure reliable cross-lingual comparative analysis.

5.4.2. Syntactic realisation of the frame elements in the Telling frame

The **Speaker** (or in its absence, the **Medium**) usually takes the position of the subject (external NP). As clearly shown in Table 5, most often the **Addressee** is expressed as an NP object (5.b.), and it can be the indirect object NP in a double-object construction (5.c.). With some verbs the **Addressee** may only be expressed as a PP, e.g. *confide* (5.a.), or, as with *tell*, may alternate with an NP object (compare 5.c. and 5.f.). The **Message** is most often realised as a subordinate clause (5.d., 6.c.), a prepositional phrase (5.e., 5.g.) or as a quote, and more rarely as an object NP (cf. Table 5). In the latter case the **Addressee** is expressed as a PP (5.a., 5.f.) or as an indirect object NP (in a double-object construction). Instead of the **Message** or alongside it, the **Topic** may be realised as a prepositional phrase (5.b.).

5.a. [A doctor]_{Speaker} must not CONFIDE [personal details]_{Message} [to a patient]_{Addressee}.

- 5.b. [They]_{Speaker} *must INFORM* [patients]_{Addressee} [about the side effects]_{Topic}.
 5.c. [My mom]_{Speaker} *TOLD* [the girls]_{Addressee} [a lot of scary stories]_{Message}.
 5.d. [We]_{Speaker} *NOTIFIED* [his cousin]_{Addressee} [that he was in hospital]_{Message}.
 5.e. [He]_{Speaker} *INFORMED* [them]_{Addressee} [of his decision]_{Message}.
 5.f. [_{Speaker}] *TELL* [that]_{Message} [to our teacher]_{Addressee}.
 5.g. [You]_{Speaker} *should NOTIFY* [them]_{Addressee} [of your wish]_{Message}.

In Bulgarian the **Addressee** is expressed as an object NP (6.b., 6.c.) or as a PP (6.a., 6.d.), depending on the particular verb. The **Message** is most often realised as a subordinate clause (6.c.) or a quote (6.a.), and only some verbs realise it as an object NP, in which case the **Addressee** is expressed as a PP (6.d.). Verbs that take an NP **Addressee** may also express the **Message** as a PP (6.e.), but such cases are missing in the Bulgarian corpus data. Instead of the **Message** or alongside it, the **Topic**, may be realised as a prepositional phrase (6.b.).

6.a. [- Днес излизате на свобода]_{Message} – КАЗВА [директорът на затвора]_{Speaker} [на затворника]_{Addressee}.

– Today you go out free – TELLS warden-DEF to convict-DEF.

‘‘Today you are being released,’’ the warden tells the convict.’

6.b. [Той]_{Speaker} *незабавно УВЕДОМЯВА* [за това]_{Topic} [програмния директор]_{Addressee}.

He immediately INFORMS about that programming-DEF director.

‘He shall immediately inform the programming director about that.’

6.c. [Лиман]_{Speaker} [го]_{Addressee} *УБЕРИ*, [че ще успеят]_{Message}.

Leman him ASSURED that they will succeed.

‘Leman assured him that they will succeed.’

6.d. [Момичето]_{Speaker} [й]_{Addressee} *ДОВЕРИ* [най-съкровените си тайни]_{Message}.

The girl to her CONFIDED deepest-DEF REFL.POSS secrets.

‘The girl confided her deepest secrets in her.’

6.e. [_{Speaker}] *УВЕДОМИ* [гу]_{Addressee} [за желанието си]_{Message}.

NOTIFY them about wish-DEF REFL.POSS.

‘You should notify them of your wish.’

6.f. [Тя]_{Speaker} [гу]_{Addressee} *УВЕДОМИ* [за напускането си]_{Message}.

She them NOTIFIED about resignation-DEF REFL.POSS.

‘She notified them about her resignation.’

The prevalent valence patterns for the verbs in the FrameNet frame **Telling** are illustrated in Table 5. The data shows considerable variation in the valence patterns in English and Bulgarian, in particular with respect to the realisation of the **Message** as an object, a clause or a quote, or with respect to allowing null-instantiations of the **Addressee**.

| Valence patterns | # | EN verbs | # | BG verbs |
|--|----|---|----|---|
| [NP.Ext] _{Speaker} [NP.Obj] _{Addressee} [Clause] _{Message} | 53 | <i>inform, advise, tell, assure, notify</i> | 32 | <i>уверявам / уверя, уведомявам / уведомя, осведомявам / осведомя</i> |
| [NP.Ext] _{Speaker} [NP.Obj] _{Addressee} [Quote] _{Message} | | | 9 | <i>уверявам / уверя, уведомявам / уведомя</i> |
| [NP.Ext] _{Speaker} [NP.Obj] _{Addressee} [PP] _{Topic} | 30 | <i>appraise, inform, advise, tell, notify</i> | 5 | <i>осведомявам / осведомя, уведомявам / уведомя</i> |
| [NP.Ext] _{Speaker} [] _{Addressee-DNI} [Clause] _{Message} | 26 | <i>advise, confide, tell, assure</i> | | |
| [NP.Ext] _{Speaker} [NP.Obj] _{Addressee} [] _{Message-DNI} | 20 | <i>inform, tell, assure, notify</i> | 3 | <i>осведомявам / осведомя, уведомявам / уведомя</i> |
| [NP.Ext] _{Speaker} [NP.Obj] _{Addressee} [PP] _{Message} | 20 | <i>inform, advise, tell, notify</i> | | |
| [NP.Ext] _{Speaker} [] _{Addressee-DNI} [NP.Obj] _{Message} | 16 | <i>advise, confide, tell</i> | | |
| [NP.Ext] _{Speaker} [PP] _{Addressee} [NP.Obj] _{Message} | 16 | <i>advise, confide, tell, notify</i> | 14 | <i>казвам / кажа, съобщавам / съобщия</i> |
| [NP.Ext] _{Speaker} [PP] _{Addressee} [Clause] _{Message} | | | 15 | <i>казвам / кажа, съобщавам / съобщия</i> |
| [NP.Ext] _{Speaker} [PP] _{Addressee} [Quote] _{Message} | | | 6 | <i>казвам / кажа</i> |

Table 5. FrameNet valence patterns of **Telling** verbs, with their frequency in the FrameNet corpus and the verbs they appear with, compared with the Bulgarian data

6. A closer look at Statement verbs

Below we take a more in-depth look at several high-frequency verbs of speech communication in Bulgarian (verbs evoking the frame **Statement**), focusing on the valence patterns and the syntactic expression of each of the considered aspect pairs as reflected in the Bulgarian annotated dataset.

We provide data and comparison among the Bulgarian verbs, on the one hand, and between each of them and its closest English counterpart, on the other. The observations reflect the distribution in the annotated datasets and may be skewed due to the limited number of instances and, possibly, selection bias. The data for English are taken from the lexical entries of the relevant verbs in FrameNet.

Tables 6a – 12a summarise the observations for the Bulgarian verbs. The members of an aspect pair are considered together. Tables 6b – 12b represent a comparison between the Bulgarian verb pairs and their correspondences in English. For a more comprehensive overall picture of the valence patterns across the discussed Bulgarian verbs, a compiled version of the monolingual Bulgarian data in Tables 6a – 12a is provided as Table 15 in the Appendix.

The pair *казвам / кажа* (Table 6a) is represented by several valence patterns involving the **Speaker** and the **Message**, with only a few instances where the **Addressee** is also expressed. Examples including realisations of the **Medium** or the **Topic** have not been found in the data. The comparison with the instances of *say* in FrameNet shows that the number of occurrences of these three frame elements is also very limited (6%, 3% and less than 2.5% of the examples for **Medium**, **Topic** and **Addressee**, respectively). While bearing out this general observation to a considerable degree, some of the verbs represented below (Tables 6a – 12a) show more prominent preference for one or another of these frame elements as compared with the rest of the predicates in the selection. For instance, *обяснявам / обясня* (explain) and *заявявам / заявя* (state) express the **Addressee** much more readily. The verbs differ with respect to the preferred syntactic expression for the individual elements as well.

The **Speaker**, as expected from the aggregated data across the semantic frames in the communication domain, is the subject, to the exception of passive sentences, where most often it remains non-overt (rarely, expressed as a prepositional phrase), and examples where the **Medium** is projected in the subject position.

Below we sum up the observations on the expression of complements. For the **Message** we observe several different valence patterns (see Section 5.3. for a synopsis of Koeva’s findings presented in Koeva 2019).

Казвам / кажа (*say, state*) and *обявявам / обявя* (*announce*) tend to express this frame element in all three possible ways: as a clause (introduced by the complementiser *че* (*that*)), as a direct quote or as an object NP. In the available dataset, the former pair, which represents the prototypical speech communication verbs, gives a slight preference to quotes (~40%) over complement clauses (33%). Both outnumber the realisation as an NP (23.5%), which most often is a non-specific expression, e.g. *нещо* (*thing, something, anything*), *нищо* (*nothing, anything*), *злуноцму* (*nonsense*), which makes a general reference to a proposition, while nominalisations and other types of nouns are not possible. By comparison, with *say*, the distribution of the valence patterns is reversed in favour of finite clauses (about half of the examples), as compared with direct quotes (about a third of the examples). Object NPs are less represented (roughly a tenth of all the instances); similarly to Bulgarian, they denote non-specific nominalised propositions expressed as *thing, anything, something, nothing*, pronouns (e.g. *what*), nominalised quantifiers (e.g. *little*) and the like.

| <i>казвам / кажа</i> (<i>say</i>) / No. sentences 51 | | | | | | | | | | | |
|--|--------|--------|----|------|----|---------------|---------------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че</i> -cl | <i>да</i> -cl | Intrg-cl | Quote | Other | Total |
| Speaker | 49 | | | | 2 | | | | | | 51 |
| Message | 1 | 12 | | | | 17 | 1 | | 20 | | 51 |
| Addressee | | | 5 | | | | | | | | 5 |

Table 6a. Valence patterns of *казвам / кажа*

| казвам / кажа No. sentences 51 / say No. sentences 371 | | | | | | | | | | | | | | | | |
|--|---------|-----|---------|----|---------|----|---------|----|---------|-----|---------|-----|---------|----|---------|-----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Quote | | Other | | Total | |
| | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | |
| Speaker | 49 | 335 | | | | | 2 | 24 | | | | | | | 51 | 359 |
| Message | 1 | 14 | 12 | 32 | | | | | 18 | 178 | 20 | 124 | | 37 | 51 | 371 |
| Topic | | | | | | 11 | | | | | | | | 2 | | 13 |
| Medium | | 11 | | | | 11 | | | | | | | | 1 | | 23 |
| Addressee | | | | | 5 | 9 | | | | | | | | | 5 | 9 |

Table 6b. Comparison of the valence patterns of *казвам / кажа* and *say*

With respect to *обявявам / обявя* the favoured expression of the **Message** is as a subordinate clause (about half of the examples), while object NPs and direct quotes are represented on a par (22.5%). Unlike *казвам / кажа*, the NPs are often nominalisations such as *решение (decision)*, *желание (wish)*, *намерение (intention)*, or nouns unrelated to verbs, e.g. *данни (data)*. *Announce* expresses the **Message** as a subordinate clause also in around half of the occurrences, but favours NP objects (a little over a quarter of the examples) over direct quotes (~15%).

| обявявам / обявя (announce) / No. sentences 49 | | | | | | | | | | | |
|--|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | 1 | | 3 | | | | | | 49 |
| Message | 2 | 11 | | | | 23 | | | 11 | 2 | 49 |
| Medium | | | 1 | | | | | | | | 1 |
| Addressee | | | 1 | | | | | | | | 1 |

Table 7a. Valence patterns of *обявявам / обявя*

| обявявам / обявя No. sentences 49 / announce No. sentences 85 | | | | | | | | | | | | | | | | |
|---|---------|----|---------|--|---------|---|---------|---|---------|--|---------|--|---------|----|---------|--|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Quote | | Other | | Total | |
| | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | |
| Speaker | 45 | 68 | | | 1 | 3 | 3 | 9 | | | | | 1 | 49 | 81 | |

| | | | | | | | | | | | | | | |
|-----------|---|---|----|----|---|---|----|----|----|----|---|---|----|----|
| Message | 2 | 8 | 11 | 22 | | | 23 | 42 | 11 | 12 | 2 | 1 | 49 | 85 |
| Medium | | 3 | | | 1 | 3 | | | | | | | 1 | 6 |
| Addressee | | | | | 1 | 7 | | | | | | 1 | 1 | 8 |

Table 7b. Comparison of the valence patterns of *объявям / объяв* and *announce*

Заявляюм / заявя (state), добавям / добавя (add), допълвам / допълня (add), отбелязвам / отбележа (note, observe) show preference for expressing the **Message** as a finite subordinate clause or as a direct quote over object NP.

More specifically, *заявляюм / заявя (state)* realises the **Message** as a subordinate *че*-clause in approximately half of the instances (47%), with individual examples of *да*-clauses and interrogative clauses, or as a quote (33%). The numbers for *state* are quite similar with respect to the proportion of finite clauses (48%) and are a little bit smaller for quotes (27%). Object NPs, although rarer, are not limited to words and expressions referring to propositions in a general way (e.g. *thing, something, nothing*), and the examples include words such as *несъгласие (disagreement), позиция (point of view)*, etc.

| | | | | | | | | | | | |
|---|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| <i>заявляюм / заявя (state, say, tell) / No. sentences 48</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | | | 3 | | | | | | 48 |
| Message | | 5 | | | | 21 | 1 | 1 | 15 | 5 | 46 |
| Medium | | | 3 | | | | | | | | 3 |
| Addressee | | | 13 | | | | | | | | 13 |

Table 8a. Valence patterns of *заявляюм / заявя*

| | | | | | | | | | | | | | | | | | | | |
|---|--------|----|--------|----|----|----|----|----|--------|----|----------|----|-------|----|-------|----|-------|----|----|
| <i>заявляюм / заявя No. Sentences 48 / state No. sentences 48</i> | | | | | | | | | | | | | | | | | | | |
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Intrg-cl | | Quote | | Other | | Total | | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | |
| Speaker | 45 | 38 | | | | | 3 | | | | | | | | | | | 48 | 38 |
| Message | | 3 | 5 | 8 | | | | | 22 | 24 | 1 | 1 | 15 | 13 | 5 | 5 | | 46 | 48 |

| | | | | | | | | | | | | | | |
|-----------|--|---|--|----|---|--|---|--|--|--|--|--|----|---|
| Medium | | 9 | | 3 | 1 | | 4 | | | | | | 3 | 4 |
| Addressee | | | | 13 | 3 | | | | | | | | 13 | 3 |

Table 8b. Comparison of the valence patterns of *заявявам* / *заявя* and *state*

The verbs *добавям* / *добавя* and its synonym *допълвам* / *допълня* (*add*) show a similar valence pattern, with prevalence of quotes (53% and 67%, respectively) over *че*-clauses (28% for both verbs). By contrast, in the English dataset we work with, its counterpart the verb *add* also selects predominantly either one or the other type of propositional complement, but the distribution is much more even (36% and 42%). NP objects constitute a small number in both languages and usually refer to words related to information content: *comments*, *details*, *information*, *words*, etc., or in the more general case, *something*, *anything* or the like.

| <i>добавям</i> / <i>добавя</i> (<i>add</i>) / No. sentences 42 | | | | | | | | | | | |
|--|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 41 | | | | | | | | | | 41 |
| Message | | 6 | | | 1 | 12 | | 1 | 23 | | 43 |
| Medium | 1 | | | | | | | | | | |
| <i>допълвам</i> / <i>допълня</i> (<i>add</i>) / No. sentences 35 | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 35 | | | | | | | | | | 35 |
| Message | | 2 | | | | 10 | | | 24 | | 36 |

Table 9a. Valence patterns of *добавям* / *добавя* / *допълвам* / *допълня*

| <i>добавям</i> / <i>добавя</i> / <i>допълвам</i> / <i>допълня</i> No. sentences 76 / <i>add</i> No. sentences 64 | | | | | | | | | | | | | | | | | | |
|--|--------|----|--------|----|----|----|----|----|--------|----|----------|----|-------|----|-------|----|-------|----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Intrg-cl | | Quote | | Other | | Total | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En |
| Speaker | 76 | 61 | | | | | | 2 | | | | | | | | 1 | 76 | 64 |
| Message | | 1 | 8 | 7 | | | 1 | | 22 | 25 | 1 | | 47 | 29 | 2 | 3 | 79 | 64 |

| | | | | | | | | | | | | | |
|-----------|---|---|--|--|---|--|--|--|--|--|--|---|---|
| Topic | | | | | 4 | | | | | | | | 4 |
| Medium | 1 | 1 | | | | | | | | | | 1 | 1 |
| Addressee | | | | | 1 | | | | | | | | 1 |

Table 9b. Comparison of the valence patterns of *добавям / добавя / допълвам / допълня* and *add*

Обяснявам / обясня (explain) selects rather equally a *че*-finite clause, an interrogative clause or a quote (~26%, 23%, 21% of the examples). This goes along with the meaning of the verb as it refers to giving details or elaborating on a certain subject and respectively on the various aspects and circumstances involved (*how, when, where, etc.*). By contrast, at least in the available data, *explain* favours finite clauses (60%) over quotes (25%) and interrogative clauses are represented by a single example. In addition, the verbs show a marked tendency to express the **Topic** (63%) over the **Message** (39%), i.e. to refer to the message by means of elaborating on its subject matter rather than on stating the message itself. For Bulgarian, this number is much smaller (10%). Judging from the examples, this may be due to differences in the construal of what the content of the **Topic** and the **Message** may be, as part of the interrogative complement clauses are annotated as **Topics**.

| обяснявам / обясня (explain) / No. sentences 49 | | | | | | | | | | | |
|---|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | | | 3 | | | | | | 48 |
| Message | 2 | 6 | | | 3 | 11 | | 10 | 9 | 2 | 43 |
| Topic | | 5 | | | | | | | | | 5 |
| Medium | 1 | | 3 | | | | | | | | 4 |
| Addressee | | | 19 | | | | | | | | 19 |

Table 10a. Valence patterns of *обяснявам / обясня*

| обяснявам / обясня No. Sentences 49 / explain No. sentences 51 | | | | | | | | | | | | | | | | | | |
|--|--------|----|--------|----|----|----|----|----|--------|----|----------|----|-------|----|-------|----|-------|----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Intrg-cl | | Quote | | Other | | Total | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En |
| Speaker | 45 | 45 | | | | 2 | 3 | 3 | | | | | | | | 1 | 48 | 51 |
| Message | 2 | | 6 | 2 | | | 3 | | 11 | 12 | 10 | 1 | 9 | 5 | 2 | | 43 | 20 |
| Topic | | 3 | 5 | 16 | | 6 | | | | | | 5 | | | | | 5 | 32 |

| | | | | | | | | | | | | |
|-----------|---|---|--|----|----|--|--|--|--|---|----|----|
| Medium | 1 | 1 | | 3 | 4 | | | | | | 4 | 5 |
| Addressee | | | | 19 | 15 | | | | | 1 | 19 | 16 |

Table 10b. Comparison of the valence patterns of *обяснявам* / *обясня* and *explain*

Отбелязвам / *отбележа* (*note, remark*) select **Messages** expressed by a *че*-clause or a quote in even proportions. As there are two suitable verbs in the English data (*note* and *remark*) we consider them both. The two verbs show distinct valence patterns: *note* has a preference for finite clauses (85%) and *remark* for quotes (61%).

| | | | | | | | | | | | |
|---|--------|--------|----|------|----|---------------|---------------|----------|-------|-------|-------|
| <i>отбелязвам</i> / <i>отбележа</i> (<i>note, observe, remark</i>) / No. sentences 49 | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че</i> -cl | <i>да</i> -cl | Intrg-cl | Quote | Other | Total |
| Speaker | 40 | | | | 5 | | | | | | 45 |
| Message | 1 | 4 | | | | 21 | | | 20 | 3 | 49 |
| Medium | 4 | | 6 | | | | | | | | 10 |
| Addressee | | | 2 | | | | | | | | 2 |

Table 11a. Valence patterns of *отбелязвам* / *отбележа*

| | | | | | | | | | | | | | | | | |
|---|---------|-------|---------|----|---------|-----|---------|----|---------|--|----------|----|---------|----|---------|-------|
| <i>отбелязвам</i> / <i>отбележа</i> No. sentences 49 / <i>note</i> No. sentences 40, <i>remark</i> No. sentences 39 | | | | | | | | | | | | | | | | |
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Intrg-cl | | Quote | | Total | |
| | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | | Bg En | |
| Speaker | 40 | 28,37 | | | 1 | 5 | 9,1 | | | | | | | | 45 | 37,39 |
| Message | 1 | | 4 | 5, | | | | 21 | 34,6 | | ,3 | 20 | ,19 | 49 | 40,31 | |
| Topic | | 1, | | 1, | | ,12 | | | | | | | | | | 2,12 |
| Medium | 4 | 5,1 | | | 6 | 5,1 | | | | | | | | 10 | 12,2 | |
| Addressee | | | | | 2 | ,3 | | | | | | | | 2 | ,3 | |

Table 11b. Comparison of the valence patterns of *отбелязвам* / *отбележа* and *note, remark*⁸

⁸ The values for *note* and *remark* are separated by a comma.

Настоявам / настоя (insist⁹) select for a **Message** that is a clause (33%), a PP (29%) or a quote (27%); while *insist* favours clauses (51% for clauses, 30% for PPs and 19% for quotes). The Bulgarian verbs are the only ones in this sample that take predominantly a subordinate clause introduced by *да* (non-factitive clauses), whose counterparts in English, strictly speaking, should be subjunctive clauses. In both languages the verbs may also take a prepositional instead of a clausal **Message**.

| <i>настоявам / настоя (insist) / No. sentences 48</i> | | | | | | | | | | | |
|---|--------|--------|----|------|----|---------------|---------------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че</i> -cl | <i>да</i> -cl | Intrg-cl | Quote | Other | Total |
| Speaker | 48 | | | | | | | | | | 48 |
| Message | | | 14 | | 5 | 4 | 12 | | 13 | | 48 |
| Addressee | | | 1 | | | | | | | | 1 |

Table 12a. Valence patterns of *настоявам / настоя*

| <i>настоявам / настоя No. sentences 48 / insist No. sentences 57</i> | | | | | | | | | | | | | | |
|--|--------|----|--------|----|----|----|----|----|--------|----|-------|----|-------|----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Quote | | Total | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En |
| Speaker | 48 | 54 | | | | 2 | | 1 | | | | | 48 | 57 |
| Message | | 2 | | | 14 | 17 | 5 | | 4+12 | 29 | 13 | 11 | 48 | 57 |
| Addressee | | | | | 1 | 3 | | | | | | | 1 | 3 |

Table 12b. Comparison of the valence patterns of *настоявам / настоя* and *insist* (The clauses introduced by *че* (4) and *да* (12) are given separately for reference.)

With respect to the **Addressee**, as noted above, some verbs, such as *заявявам / заявя* and especially *обяснявам / обясня* favour the expression of the (intended) recipient of the message, as their semantics imply the presence of an **Addressee** to whom the content (explanation or statement) is directed. The same goes for *announce* and *explain* in English, and to a lesser extent to both *казвам / кажа* and *say*. The remaining verbs only select for an overt **Addressee** occasionally.

⁹ We discuss the sense corresponding to the FrameNet lexical unit evoking the **Statement** frame and defined as ‘demand or state forcefully, without accepting refusal or contradiction’. The relevant sense is not described in WordNet.

The small number of examples for **Medium** and **Topic** do not allow us to make reliable conclusions. The extension of the datasets would help corroborate or correct the observations.

7. The Addressee as a core frame element in the frame Telling

As already discussed above, the frame **Telling** inherits from **Statement** and elaborates on it by moving the focus of the situation from the **Speaker** and the **Message** onto the **Addressee** receiving the message. This results in the promotion of the **Addressee** to a core status and consequently, its more often than not, overt expression.

In both English and Bulgarian, the frame **Telling** is evoked by two groups of verbs with respect to the realisation of the **Addressee**: (1) verbs that require the **Addressee** to be realised as a direct object NP, and (2) verbs that realise the **Addressee** as an indirect object (as a prepositional phrase, or as a dative pronominal clitic). Tables 13a and 14a show the valence patterns of example verbs from each group; Tables 13b and 14b present a comparison between the Bulgarian verb pairs and their correspondences in English.

Verbs such as *inform*, *notify*, *advise*, *assure* in English, and *уведомявам* / *уведомя* (*notify*, *inform*), *известявам* / *известя* (*notify*), *информирам* (*inform*), *осведомявам* / *осведомя* (*inform*) in Bulgarian, belong to the first group for which the **Addressee** is realised as an NP object and the **Message** is expressed as an indirect object, a clause or a quote.

| уведомявам / уведомя (inform) / No. sentences 16 | | | | | | | | | | | |
|--|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | ce-cl | da-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 16 | | | | | | | | | | 16 |
| Message | | | | | 5 | 4 | | 1 | 3 | | 13 |
| Topic | | | 4 | | | | | | | | 4 |
| Addressee | | | 15 | | 1 | | | | | | 16 |

Table 13a. Valence patterns of *уведомявам* / *уведомя*

| уведомявам / уведомя No. sentences 16 / inform No. sentences 39 | | | | | | | | | | | | | | |
|---|--------|----|--------|----|----|----|----|----|--------|----|-------|----|-------|----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Quote | | Total | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En |
| Speaker | 16 | 39 | | | | | | | | | | | 16 | 39 |
| Message | | | | | | 9 | 5 | 3 | 5 | 19 | 3 | 6 | 13 | 37 |

| | | | | | | | | | | | | | | |
|-----------|--|--|----|----|---|---|---|---|--|--|--|--|----|----|
| Topic | | | | | 4 | 4 | | | | | | | 4 | 4 |
| Addressee | | | 15 | 37 | | | 1 | 2 | | | | | 16 | 39 |

Table 13b. Comparison of the valence patterns of *уведомявам / уведомя* and *inform*

The second group includes verbs such as *confide* in English and *казвам / кажа* (*tell*), *съобщавам / съобщя* (*inform*), *доверявам / доверя* (*confide*) in Bulgarian, for which the **Addressee** is also compulsory but assumes the position of the indirect object as the receiver to whom the message is directed; the **Message** is realised as a direct object, a clause or a quote.

In Bulgarian the aspectual pair *казвам / кажа* (among others) is ambiguous and can correspond to a number of synsets, out of which we are interested in {state, say, tell}, ‘express in words’ (evoking the frame **Statement**) and {tell}, ‘let something be known’ (evoking the frame **Telling**). While in English the verb *tell* can have as a direct object either the **Message** or the **Addressee**, or even express them in a double object construction (5.c.), the verb pair *казвам / кажа* always realises the **Message** in the direct object position and the **Addressee** as an indirect object.

| казвам / кажа (tell) / No. sentences 32 (frame Telling) | | | | | | | | | | | |
|---|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 32 | | | | | | | | | | 32 |
| Message | | 11 | | | | 15 | | | 6 | | 32 |
| Topic | | | 6 | | | | | | | | 6 |
| Addressee | | | 32 | | | | | | | | 32 |

Table 14a. Valence patterns of *казвам / кажа* (frame **Telling**)

| казвам / кажа No. sentences 32 / tell No. sentences 104 (frame Telling) | | | | | | | | | | | | | | | | |
|---|--------|----|--------|----|----|----|----|----|--------|----|-------|----|-------|----|-------|-----|
| | NP.Ext | | NP.Obj | | PP | | NI | | Fin-cl | | Quote | | Other | | Total | |
| | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En | Bg | En |
| Speaker | 32 | 90 | | | 9 | | 14 | | | | | | | | 32 | 104 |
| Message | | | 11 | 11 | 9 | | 33 | 15 | 35 | 6 | 6 | | 10 | | 32 | 104 |
| Topic | | | | | 6 | 31 | | | | | | | | | 6 | 31 |
| Addressee | | | 59 | 32 | 3 | | 42 | | | | | | | | 32 | 104 |

Table 14b. Comparison of the valence patterns of *казвам / кажа* and *tell* (frame **Telling**)

This is a notable syntactic difference between the corresponding verbs in the two languages, which results in a language specificity of the syntactic expression that needs to be accounted for in the description of the **Telling** frame in English and Bulgarian.

8. Conclusions and future work

The analysis of the conceptual and syntactic properties of English and Bulgarian verbs as attested in the corpora of annotated examples, have helped confirm the applicability of the description provided in the FrameNet frames and annotated dataset to the analysis of Bulgarian verbs by employing the aspects of the semantic and syntactic representation that are relatively language-independent and transferrable cross-linguistically and making the necessary adjustments, where needed.

A study based on corpus analysis and statistical observations on the frequency of valence patterns could provide more reliable evidence for the behaviour of verbs, in particular in view of cross-linguistic studies. Moreover, this will confirm the validity of the cross-linguistic analysis and the universality of semantic and syntactic features.

For Bulgarian and English we have demonstrated substantial correspondence both in terms of the valence patterns and the syntactic categories and grammatical functions whereby frame elements are expressed. A more comprehensive study involving other languages (Slavic and Balkan languages, in particular) may provide a solid theoretical and methodological foundation for comparative/contrastive research into syntax and semantics.

Acknowledgments: This study is carried out as part of the scientific programme under the project *Enriching the Semantic Network WordNet with Conceptual Frames* funded by the Bulgarian National Science Fund (Grant Agreement No. KP-06-N50/1 of 2020).

References

- Baker 2008: Baker, C. F. FrameNet: Present and Future. – In: Webster, J., N. Ide, A. C. Fang (Eds.). *The First International Conference on Global Interoperability for Language Resources*. Hong Kong: City University.
- Baker et al. 1998: Baker, C. F., C. J. Fillmore, J. B. Lowe. The Berkeley FrameNet project. – In: *COLING-ACL '98: Proceedings of the Conference*. Montreal, Canada, pp. 86 – 90.
- Baker, Fellbaum 2009: Baker, C. F., C. Fellbaum. WordNet and FrameNet as Complementary Resources for Annotation. – In: *Proceedings of the Third Linguistic Annotation Workshop (ACL-IJCNLP '09)*. Stroudsburg: Association for Computational Linguistics, pp. 125 – 129.
- Boas 2010: Boas, H. The syntax-lexicon continuum in Construction Grammar: A case study of English communication verbs. – *Belgian Journal of Linguistics*, 24, pp. 54 – 82.

- Burchardt et al. 2005: Burchardt, A., K. Erk, A. Frank. A WordNet detour to FrameNet. – In: *Sprachtechnologie, mobile Kommunikation und linguistische Ressourcen*, vol. 8 (*Computer Studies in Language and Speech*). Frankfurt: Lang.
- Das et al. 2014: Das, D., D. Chen, A. F. T. Martins, N. Schneider, N. A. Smith. Frame-Semantic Parsing. – *Computational Linguistics*, 40(1), pp. 9 – 56.
- de Lacalle et al. 2014: de Lacalle, M. L., E. Laparra, G. Rigau. Predicate Matrix: extending SemLink through WordNet mappings. – In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik: European Language Resources Association (ELRA), pp. 903 – 909.
- Fabio et al. 2019: Fabio, A. Di, S. Conia, R. Navigli. VerbAtlas: a Novel Large-Scale Verbal Semantic Resource and Its Application to Semantic Role Labeling. – In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong: Association for Computational Linguistics, pp. 627 – 637.
- Fellbaum 1998a: Fellbaum, C. A Semantic Network of English Verbs. In: Fellbaum, C. (Ed.). *WordNet: an electronic lexical database*. Cambridge, MA: MIT Press, pp. 69 – 104.
- Fellbaum 1998b: Fellbaum, C. (Ed.). *WordNet: an electronic lexical database*. Cambridge, MA: MIT Press.
- Fellbaum 1999: Fellbaum, C. The Organization of Verbs and Verb Concepts in a Semantic Net. – In: Saint-Dizier, P. (Ed.). *Predicative Forms in Natural Language and in Lexical Knowledge Bases. Text, Speech and Language Technology*, vol. 6. Dordrecht: Springer.
- Fillmore 1982: Fillmore, C. Frame semantics. – In: *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co.
- Jackendoff 1990: Jackendoff, R. *Semantic structures*. MIT Press.
- Kipper et al. 2008: Kipper, K., A. Korhonen, N. Ryant, M. Palmer. A large-scale classification of English verbs. – *Language resources and evaluation. Communications*, 42(1), pp. 21 – 40.
- Kipper-Schuler 2005: Kipper-Schuler, K. *VerbNet: A broad-coverage, comprehensive verb lexicon. PhD Thesis*. Philadelphia, PA: Computer & Information Science Dept., University of Pennsylvania.
- Коева 2010: Коева, С. *Българският ФреймНет*. София: Институт за български език Проф. Любомир Андрейчин. (Коева, S. *Balgarskiyat FreymNet*. Sofia: Institute for Bulgarian Language.)
- Коева 2014: Коева, С. WordNet и БулНет. – В: Коева, С. (съст.). *Езикови ресурси и технологии за българския език*. София: Академично издателство „Проф. Марин Дринов“, с. 154 – 173. (Коева, S. *WordNet i BulNet*. – In: Коева, S. (sast.). *Ezikovi resursi i tehnologii za balgarskiya ezik*. Sofia: Prof. Marin Drinov Publishing House of BAS, pp. 154 – 173.)
- Коева 2019: Коева, С. Комплементите в български. – В: Сборник с доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“. София: Издателство на БАН „Проф. Марин Дринов“, с. 57 – 68. (Коева, S. *Komplementite v balgarski*. – In: *Proceedings of the International Annual Conference of the Institute of Bulgarian Language*

- Prof. Lubomir Andreychin*. Sofia: Prof. Marin Drinov Publishing House of BAS, pp. 57 – 68.)
- Koeva 2020: Koeva, S. Semantic relations and conceptual frames. – In: Koeva, S. (Ed.). *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Institute for Bulgarian Language, pp. 7 – 20.
- Koeva 2021: Koeva, S. The Bulgarian WordNet: Structure and specific features. – *Papers of Bulgarian Academy of Sciences*, 8(1), pp. 47 – 70. <https://www.papersofbas.eu/images/Papers_2021-1/Koeva_optimized.pdf> [18.01.2024]
- Koeva et al. 2006: Koeva, S., S. Leseva, M. Todorova. Bulgarian sense tagged corpus. – In: *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*. Genoa: European Language Resources Association (ELRA), pp. 79 – 86.
- Koeva et al. 2011: Koeva, S., S. Leseva, B. Rizov, E. Tarpomanova, T. Dimitrova, H. Kukova, M. Todorova. Design and development of the Bulgarian sense-annotated corpus. – In: Carry Pastor, M. L., M. A. Candel Mora (Eds.). *Information and communications technologies: present and future in corpus analysis: Proceedings of the III International Congress of Corpus Linguistics*. Valencia: Valencia Polytechnic University, pp. 143 – 150.
- Koeva et al. 2012a: Koeva, S., B. Rizov, E. Tarpomanova, T. Dimitrova, R. Dekova, I. Stoyanova, S. Leseva, H. Kukova, A. Genov. Bulgarian-English sentence- and clause-aligned corpus. – In: *Proceedings of the Second Workshop on Annotation of Corpora for Research in the Humanities (ACRH-2)*. Lisboa: Colibri, pp. 51 – 62.
- Koeva et al. 2012b: Koeva, S., I. Stoyanova, S. Leseva, R. Dekova, T. Dimitrova, E. Tarpomanova. The Bulgarian National Corpus: Theory and Practice in Corpus Design. – *Journal of Language Modelling*, (1), pp. 65 – 110.
- Koeva, Dekova 2008: Koeva, S., R. Dekova. Bulgarian FrameNet. – In: Tadic, M., M. Vulchanova, S. Koeva (Eds.). *Proceedings from The Sixth International Conference Formal Approaches to South Slavic and Balkan Languages*. Dubrovnik: Croatian Language Technologies Society, Faculty of Humanities and Social Science pp. 59 – 67.
- Koeva, Doychev 2022: Koeva, S., E. Doychev. Ontology supported frame classification. – In: *Proceedings of the Fifth International Conference on Computational Linguistics in Bulgaria*. Sofia: Department of Computational Linguistics, Institute for Bulgarian Language – Bulgarian Academy of Sciences, pp. 203 – 214. <<https://aclanthology.org/2022.clib-1.23>> [18.01.2024]
- Landes et al. 1998: Landes, S., C. Leacock, R. Tengi. Building Semantic Concordances. – In: C. Fellbaum (Ed.). *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press, pp. 199 – 216.
- Laparra, Rigau 2009: Laparra, E., G. Rigau. Integrating WordNet and FrameNet using a knowledge-based Word Sense Disambiguation algorithm. – In: *Proceedings of Recent Advances in Natural Language Processing (RANLP09), Borovets, Bulgaria*, pp. 208 – 213.
- Leseva, Stoyanova 2019: Leseva, S., I. Stoyanova. Enhancing Conceptual Description through Resource Linking and Exploration of Semantic Relations. – In: *Proceedings of the Tenth Global Wordnet Conference, 23 – 27 July 2019, Wrocław, Poland*. Oficyna Wydawnicza Politechniki Wrocławskiej, pp. 280 – 289. <<https://aclanthology.org/2019.gwc-1.36.pdf>> [18.01.2024]

- Leseva, Stoyanova 2020: Leseva, S., I. Stoyanova. Towards a Conceptual Description of Verbs. – *Cybernetics and Information Technologies*, 20, 4 (Nov 2020), pp. 108 – 124.
- Levin 1993: Levin, B. *English verb classes and alternations: A preliminary investigation*. Chicago & London: The University of Chicago Press.
- Levin et al. 1997: Levin, B., G. Song, B. T. S. Atkins. Making sense of corpus data: a case study of verbs of sound. – *International Journal of Corpus Linguistics*, 2, pp. 23 – 64.
- Litkowski 2014: Litkowski, K. The FrameNet Frame Element Taxonomy. <<http://www.clres.com/online-papers/FETaxonomy.pdf>> [18.01.2024]
- Lopatkova et al. 2016: Lopatková, M., V. Kettnerová, E. Bejček, A. Vernerová, Z. Žabokrtský, Valenční slovník českých sloves VALLEX, Nakladatelství Karolinum, Praha, 2016.
- Miller 1995: Miller, G. A. WordNet: A lexical database for English. – *Communications of the ACM*, 38(11), pp. 39 – 41.
- Miller et al. 1993a: Miller, G., R. Beckwith, C. Fellbaum, D. Gross, K. Miller. Introduction to WordNet: an On-line Lexical Database. – In: *Five Papers on WordNet*. Princeton, NJ: Princeton University.
- Miller et al. 1993b: Miller, G. A., C. Leacock, R. Teng, R. T. Bunker. A Semantic Concordance. – In: *Human Language Technology: Proceedings of a Workshop Held at Plainsboro, New Jersey, March 21-24, 1993*. <<https://aclanthology.org/H93-1061>> [18.01.2024]
- Miller et al. 1994: Miller, G. A., M. Chodorow, S. Landes, C. Leacock, R. G. Thomas. Using a Semantic Concordance for Sense Identification. – In: *Human Language Technology: Proceedings of a Workshop held at Plainsboro, New Jersey, March 8-11, 1994*. <<https://aclanthology.org/H94-1046>> [18.01.2024]
- Navigli, Ponzetto 2010: Navigli, R., S. P. Ponzetto. BabelNet: Building a Very Large Multilingual Semantic Network. – In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Uppsala: Association for Computational Linguistics, pp. 216 – 225.
- Palmer 2009: Palmer, M. Semlink: linking PropBank, VerbNet and FrameNet. – In: *Proceedings of the 5th International Conference on Generative Approaches to the Lexicon. Sept. 2009, Pisa, Italy: GenLex-09*. Pisa, pp. 9 – 15.
- Palmer et al. 2014: Palmer, M., C. Bonial, D. McCarthy. SemLink+: FrameNet, VerbNet and event ontologies. – In: *Proceedings of Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*. Baltimore, Maryland USA: Association for Computational Linguistics, pp. 13 – 17.
- Petruck 2015: Petruck, M. *The Components of FrameNet. Tutorial at NAACL-HLT 2015*. <<http://naacl.org/naacl-hlt-2015/tutorial-framenet-data/FNComponentsMRLP.pdf>> [18.01.2024]
- Petruck 2019: Petruck, M. Meaning Representation of Null Instantiated Semantic Roles in FrameNet. – In: *Proceedings of the First International Workshop on Designing Meaning Representations*. Association for Computational Linguistics, pp. 121 – 127.
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: extended theory and practice*.

- Berkeley: International Computer Science Institute. <<https://framenet2.icsi.berkeley.edu/docs/r1.7/book.pdf>> [18.01.2024]
- Schneider et al. 2012: Schneider, N., B. Mohit, K. Oflazer, N. A. Smith. Coarse Lexical Semantic Annotation with Supersenses: An Arabic Case Study. – In: *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Association for Computational Linguistics, pp. 253 – 258.
- Shi, Mihalcea 2005: Shi, L., R. Mihalcea. Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. – In: Gelbukh, A. (Ed.). *Computational Linguistics and Intelligent Text Processing. CICLing 2005. Lecture Notes in Computer Science*, vol. 3406. Berlin, Heidelberg: Springer.
- Tonelli, Pighin 2009: Tonelli, S., D. Pighin. New Features for Framenet – Wordnet Mapping. – In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL'09)*. Boulder, USA.
- Urban, Ruppenhofer 2001: Urban, M., J. Ruppenhofer. Shouting and Screaming: Manner and Noise Verbs in Communication. – *Literary and Linguistic Computing*, 16, pp. 77 – 97.
- Uresova et al. 2020a: Urešová, Z., E. Fučíková, E. Hajičová, J. Hajič. SynSemClass Linked Lexicon: Mapping Synonymy between Languages. – In: *Proceedings of the Globalex Workshop on Linked Lexicography, Language Resources and Evaluation Conference (LREC 2020), Marseille, 11–16 May 2020*. European Language Resources Association, pp. 10 – 19.
- Uresova et al. 2020b: Urešová, Z., E. Fučíková, E. Hajičová, J. Hajič. Syntactic-Semantic Classes of Context-Sensitive Synonyms Based on a Bilingual Corpus. – In: Vetulani, Z., P. Paroubek, M Kubis (Eds.). *Human Language Technology. Challenges for Computer Science and Linguistics*. Springer International Publishing, pp. 242 – 255.
- Wierzbicka 1987: Wierzbicka, A. *English Speech Act Verbs: A Semantic Dictionary*. Sydney Academic Press.

Appendix

| казвам / кажа (say) / No. sentences 51 | | | | | | | | | | | |
|---|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 49 | | | | 2 | | | | | | 51 |
| Message | 1 | 12 | | | | 17 | 1 | | 20 | | 51 |
| Addressee | | | 5 | | | | | | | | 5 |
| обявявам / обявя (announce) / No. sentences 9 | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | да-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | 1 | | 3 | | | | | | 49 |
| Message | 2 | 11 | | | | 23 | | | 11 | 2 | 49 |

| | | | | | | | | | | | |
|---|--------|--------|----|------|----|--------------|--------------|----------|-------|-------|-------|
| Medium | | | 1 | | | | | | | | 1 |
| Addressee | | | 1 | | | | | | | | 1 |
| <i>заявявам / заявя (state, say, tell) / No. sentences 48</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че-cl</i> | <i>џа-cl</i> | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | | | 3 | | | | | | 48 |
| Message | | 5 | | | | 21 | 1 | 1 | 15 | 5 | 46 |
| Medium | | | 3 | | | | | | | | 3 |
| Addressee | | | 13 | | | | | | | | 13 |
| <i>добавям / добавя (add) / No. sentences 42</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че-cl</i> | <i>џа-cl</i> | Intrg-cl | Quote | Other | Total |
| Speaker | 41 | | | | | | | | | | 41 |
| Message | | 6 | | | 1 | 12 | | 1 | 23 | | 43 |
| Medium | 1 | | | | | | | | | | |
| <i>допълвам / допълня (add) / No. sentences 35</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че-cl</i> | <i>џа-cl</i> | Intrg-cl | Quote | Other | Total |
| Speaker | 35 | | | | | | | | | | 35 |
| Message | | 2 | | | | 10 | | | 24 | | 36 |
| <i>обяснявам / обясня (explain) / No. sentences 49</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че-cl</i> | <i>џа-cl</i> | Intrg-cl | Quote | Other | Total |
| Speaker | 45 | | | | 3 | | | | | | 48 |
| Message | 2 | 6 | | | 3 | 11 | | 10 | 9 | 2 | 43 |
| Topic | | 5 | | | | | | | | | 5 |
| Medium | 1 | | 3 | | | | | | | | 4 |
| Addressee | | | 19 | | | | | | | | 19 |
| <i>отбелязвам / отбележа (note, observe, remark) / No. sentences 49</i> | | | | | | | | | | | |
| | NP.Ext | NP.Obj | PP | AdvP | NI | <i>че-cl</i> | <i>џа-cl</i> | Intrg-cl | Quote | Other | Total |
| Speaker | 40 | | | | 5 | | | | | | 45 |
| Message | 1 | 4 | | | | 21 | | | 20 | 3 | 49 |
| Medium | 4 | | 6 | | | | | | | | 10 |
| Addressee | | | 2 | | | | | | | | 2 |

| настоявам / настоя (<i>insist</i>) / No. sentence 48 | | | | | | | | | | | |
|--|--------|--------|----|------|----|-------|-------|----------|-------|-------|-------|
| | NP.Ext | NP.Obj | PP | AdvP | NI | че-cl | џа-cl | Intrg-cl | Quote | Other | Total |
| Speaker | 48 | | | | | | | | | | 48 |
| Message | | | 14 | | 5 | 4 | 12 | | 13 | | 48 |
| Addressee | | | 1 | | | | | | | | 1 |

Table 15. Aggregated data for the valence patterns across the discussed Bulgarian verbs of the frame **Statement** (a compiled version of the data in Tables 6a – 12a)

СЕМАНТИЧНИ ФРЕЙМОВЕ ПРИ ГЛАГОЛИТЕ ЗА КОМУНИКАЦИЯ: КОРПУСНО ИЗСЛЕДВАНЕ

Светлозара Лесева, Ивелина Стоянова

Институт за български език „Проф. Любомир Андрейчин“, Българска
академия на науките

zarka@dcl.bas.bg, iva@dcl.bas.bg

Резюме. В студията се изследват свойствата на глаголите за комуникация с фокус върху предикатите, принадлежащи към няколко основни семантични фрейма за комуникация във Фреймнет. Анализът е съсредоточен върху семантичното описание на глаголите и валентните модели, представящи комбинаториката и синтактичната реализация на фреймовите елементи (елементи на семантичното описание, които в своята конфигурация представят основната семантика на даден фрейм), описващи тези глаголи в английски и български. За целите на изследването се използват два големи типа семантични ресурси: а) Принстънският уърднет (Fellbaum 1998b) и Българският уърднет, Булнет (Коева 2021), от една страна, и б) Фреймнет (Baker et al. 1998), от друга.

В изследването се дискутира общата организация на глаголната лексика, служеща за изразяване на (речево) общуване. Семантичното поле е йерархично организирано в система от семантични фреймове, които наследяват и по различен начин детайлизират или специализират основните елементи на прототипния семантичен фрейм **Комуникация (Communication)**. Въз основа на корпусни данни, извлечени от семантично аотираните корпуси Семкор (Miller et al. 1993b) за английски и Булсемкор (Коева et al. 2006) за български, в рамките на разработката на анализ се подлагат семантичните свойства и синтактичната реализация на групи глаголи, принадлежащи към няколко представителни семантични фрейма за комуникация. В хода на изследва-

нето се дискутират и онагледяват универсалните и езиково специфичните аспекти на семантичното и синтактичното описание и преносимостта им между езиците.

Извършените наблюдения относно валентните модели и синтактичната реализация на основните (ядрените) фреймови елементи ще послужат за проверка на достоверността на фреймовете, приписани на изследваните глаголи. Наред с това се извеждат и основни прилики и разлики както между глаголите в рамките на българския език, така и между тях и съответствията им в английски.

Ключови думи: глаголи за комуникация, Уърднет, Фреймнет, валентни модели, корпуси

Svetlozara Leseva
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113
Bulgaria
<https://orcid.org/0000-0001-8198-4555>

Ivelina Stoyanova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52, Shipchenski prohod Blvd., Bl. 17
Sofia 1113
Bulgaria
<https://orcid.org/0000-0003-3771-435X>

ЗА ПРИРОДАТА НА ПРИТЕЖАТЕЛИТЕ

Цветана Димитрова

Институт за български език „Проф. Любомир Андрейчин“
Българска академия на науките

cvetana@dcl.bas.bg

Резюме. Студията представя преглед на начините за изразяване на функцията притежател чрез съществителни имена от определени семантични класове, които отговарят на изискванията за реализация на съответните фреймови елементи във Фреймнет (FrameNet). Допълнително се разглеждат реализациите на фреймови елементи, които изразяват различни притежатели, при основни глаголи за трансфер на притежание в ограничен брой корпусни примери, в които експлицитно се реализират, освен фреймовия елемент (обект на трансфера), и двата фреймови елемента, между които протича трансферът, както и неядрени фреймови елементи. Въз основа на реализациите на фреймовите елементи в корпусните примери се дефинират и множества от (подходящи) съществителни за реализацията им.

Ключови думи: *притежание, Уърднет, Фреймнет*

1. Притежател: дефиниция и характеристики

Притежанието е бикултурно понятие с неясни граници (Seiler 1983: 4), което дефинира концептуално пространство с различни подобласти въз основа на взаимодействието на участващите в него елементи. В уводната статия на сборника *Dimensions of Possession* М. Херслунд и И. Барон (Herslund, Baron 2001: 1) посочват, че притежателността е понятие, а не лингвистична конструкция, така че проблемът за дефинирането ѝ произтича от разнообразните средства и начини за кодирането (чрез т.нар. притежателни конструкции) на широк диапазон от семантични отношения, повечето от които са несводими към чистото притежание в собствен смисъл. Всъщност става дума за отношение на свързаност между две същности – дефинирани най-общо като

* Цветана Димитрова. За природата на притежателите. – В: Св. Коева (ред.) *Семантични студии. Синтактично и семантично описание на български глаголи*. София: Издателство на БАН „Проф. Марин Дринов“, с. 117 – 133.
<https://doi.org/10.7546/SemanticStudies2024.05>

притежател и *притежавано*, които могат да имат различни характеристики, а оттук и взаимоотношенията между тях могат да бъдат описвани в различни граници.

В езиковедската литература са известни най-общо две основни категоризации на притежанието – постоянно или временно и отчуждаемо или неотчуждаемо (Heine 1997: 34 – 41; Stassen 2009: 15 – 18). Според Л. Стасен само отчуждаемата посесивност е прототипична, тоест само при нея притежателят (който е задължително одушевен или поне има способността да изявява своята воля) може да бъде квалифициран като собственик на притежаваното, както и да прекрати притежателното отношение (Stassen 2009: 17). Оттук авторът дефинира и две семантични характеристики – [\pm контрол] и [\pm постоянен контакт] – чрез които може да се различи отчуждаемото от неотчуждаемото притежание.

Понятие, огледално на понятието за собственост, е понятието за принадлежност, което може да се моделира и като вид постоянно локативно отношение, често кодирано в европейските езици чрез формулата *Y e na X* (която за Е. Бенвенист е огледална на *X има Y* (Benveniste 1966: 197).

Неотчуждаемото постоянно притежание най-често се свързва с отношението на част и цяло – постоянни притежания са например частите от (нечие) тяло. Това са същности като крак – на човек или животно, но и на предмет като маса или стол – които представляват елементи, задължително обвързани с цялото, т.е. с друг елемент, от който са част и който е техен притежател, но от който не могат да бъдат отделени, без да загубят съществена част от природата си. За сравнение и клонът на дървото, и крилото на самолета могат да бъдат отделени, но в дефиницията им винаги присъства връзката с цялото, от което са част. Неотчуждаеми притежания са и някои абстрактни същности като черти на характера (на човек), постоянни характеристики на определени обекти (като цвета на козината на животно, материала, от който е изработен даден предмет, както и предназначението на някои предмети) или дори признаци или характеристики на определени абстрактни понятия (например нерешимостта на определен проблем или доказуемостта на хипотеза).

В реда на постоянните притежания се класифицират често и понятията за роднински (и някои други близки) отношения като син, дъщеря, майка, баща, внук, които също изискват наличие на друг обект, с когото задължително трябва да са свързани (*X* е майка само по отношение на *Y*). Всъщност обаче тези понятия, както и редица други понятия за лични (и професионални) взаимоотношения по-скоро назовават функции, а не конкретен референт. Именно затова те не се представят експлицитно чрез релация на притежание в ресурсите, които ще бъдат разгледани.

Временното притежание може да се отнася за преместими и прехвърляеми притежания като имоти, дрехи, движими вещи (които същевременно са и отчуждаеми притежания), но също така и за някои временни състоя-

ния – заболявания, емоции и др. (които са по-скоро неотчуждаеми, макар и временни). Те също не са пряко свързани с релацията за притежание в разглежданите ресурси.

В езиковите конструкции притежателят може да се реализира както в рамките на именната фраза като зависим елемент от притежаваното, което най-често е опора на фразата (и е изразено със съществително име — срв. *къщата на Иван, Ивановата къща*), така и на равнището на изречението (например конструкции с външен притежател – срв. *Разресах ѝ косата, Видях ѝ лицето* (повече за тези конструкции вж. у Payne, Varshi 1999). Освен това притежанието може да се прехвърля, като това значение се отразява от предикатите за трансфер на притежание (които са разгледани и у Levin 1993; Rappaport Hovav, Levin 2008).

В настоящата студия разглеждаме по какъв начин тези същности са отразени в два езикови ресурса – Принстънския уърднет (и паралелния му Български уърднет) – в Част 2., и във Фреймнет – в Част 3. В Част 4. въз основа на примери¹ за употребите на предикати за четири семантични фрейма от Фреймнет, изразяващи трансфер на притежание, ще бъде разгледана реализацията на фреймовия елемент **Реципиент (Recipient)**, както и на други фреймови елементи, които изпълняват реципиентна функция. Наблюденията върху примерите позволяват да се направи извод за семантичния клас на съществителните, които биха удовлетворили изискванията за съчетаемост при синтактичната им реализация (Коева 2021b: 182). Методология за приложение на данни от корпуси за валидиране на селективните ограничения, зададени в семантичното описание на глаголите във Фреймнет и Уърднет, е дискутирана и у (Лесева, Стоянова / Leseva, Stoyanova 2020).

2. Изразяване на притежание в Уърднет

Уърднет (Принстънският уърднет) е лексикално-семантична мрежа на понятия (в английския език), които са представени под формата на синонимни множества (синсети), свързани помежду си с различни семантични, деривационни и други връзки (релации) (Miller et al. 1993). Към Принстънския уърднет са съотнесени подобни лексикално-семантични мрежи за други езици, една от които е Българският уърднет (Булнет) (Коева 2021a). Семантичното описание на съществителните и на глаголите в Уърднет включва и категоризирането им в общи семантични класове въз основа на приписани семантични примитиви – глаголите са класифицирани в 15, а имената – в 25 семантични класа (Miller et al. 1993; Fellbaum et al. 2009). Референцията към отношението на притежание е отразена съответно чрез семантичните класо-

¹ Примерите са извлечени от българския корпус, достъпен чрез приложението СкечЕнджин (<https://www.sketchengine.eu/>).

ве съществителни за притежание (noun.possession) и глаголи за притежание (verb.possession).

Като съществителни за притежание (noun.possession) са класифицирани имената, които са част от хиперонимо-хипонимното дърво на синсета {собственост; имущество; притежание} ‘всеки материален или нематериален предмет, който принадлежи на някого’, който е хипоним на {отношение; връзка; релация} ‘абстрактна представа за свойство, признак, споделен от два (или повече) обекта или части от цяло’. Под това понятие на различни „клони“ в хиперонимо-хипонимното дърво се нареждат понятия за активи, материални активи, богатство, имане, имущество, сума, ресурс, капиталовложение, капитал и под., повечето от които могат да се определят като рефериращи към отчуждаеми притежания.

Неотчуждаемите същности се категоризират в няколко различни други класа. Понятията, които са класифицирани със семантичния клас част на тялото (noun.body), са хипоними на {телесна част; част на тяло; част от тяло} ‘общо наименование на всяка отделна структурно и/или функционално обособена част на даден организъм – орган, крайник и т. н.’. Например {ръка} ‘горният (гръдният) крайник на човешкото тяло, включващ раменния пояс с рамото, мишницата, лакътя, предмишницата и китката с дланта и пръстите’ е класифицирана със семантичния клас част на тялото на човешко същество (noun.body), докато {лапа} ‘стъпало с нокти и удебелени възглавнички, предпазващи ходилото от нараняване, при четирикраките животни’ е класифицирана със семантичния клас животно (noun.animal) като част от тялото на животно.

Като растение (с примитив noun.plant) са класифицирани понятия, които са части от растение и които са хипоними на {растителна структура} ‘в ботаниката – всяка от различни видове структури, образуващи се в процеса на зараждане и развитие на растенията и гъбите, които притежават специфично устройство и функции, свързани с определени жизнени процеси’ – например {пъпка; цветна пъпка} ‘част от растение, която напомня пъпка и от която се развива цветът’. От своя страна, всички тези понятия са част от хиперонимо-хипонимното дърво на {естествен обект; природен обект} ‘естествено съществуващ обект, възникнал в резултат на непредизвикани, естествено протичащи физически, химически, биохимически и т.н. процеси’, който е хипоним на {цялост; цяло; едно цяло} ‘съвкупност от части, елементи, които съставят или се възприемат като единен обект или нещо хомогенно, неразчленено, неразделено’.

Като артефакт (noun.artifact) са определени понятия, които са части от корпуса на неодоушевен обект – например {крило; самолетно крило; крило на самолет} ‘всяка от двете хоризонтални носещи повърхности на летателен апарат, със специална форма и профил в зависимост от предназначението си, разположени от двете страни на корпуса на летателното устройство’. Те са част от хиперонимо-хипонимното дърво на {произведение; артефакт; артефакт} ‘произведен или изработен предмет, резултат на човешката дейност’, което също е хипоним на {цялост; цяло; едно цяло}.

Отношенията между понятията за дете и за родител не са хиперонимо-хипонимни, а са концептуализирани като противоположности – {дете; рожба; чедо} ‘човешко потомство, син или дъщеря на каквато и да е възраст, по отношение на своите родители’ е антоним на {родител} ‘човек, който е биологичен родител или поема функциите на такъв по отглеждането и възпитанието на дете в качеството си на приеман родител, осиновител, настойник и др.’. Тук и двете понятия са мероними, т.е. са части на понятието {семейство} ‘най-малката, основна социална единица, състояща се от двойка съпрузи и техните деца’. През антонимната двойка са свързани и понятията {майка} ‘родител от женски пол по отношение на своите деца’ и {баща} ‘родител от мъжки пол по отношение на своите деца’, които са хипоними на {родител}. Понятието {женска рожба} ‘дете от женски пол по отношение на своите родители’ с хипоним {дъщеря} ‘човешка рожба, наследник от женски пол’ е антоним на понятието {мъжка рожба} ‘дете от мъжки пол по отношение на своите родители’ с хипоним {син} ‘човешка рожба, наследник от мъжки пол’.

Ограниченията за семантичния клас на обекта, който влиза във функцията на притежавано в отношението за притежание, са сравнително малко, оттам според нас и ограниченията за характера на притежателя (или обекта, с който тези същности влизат в отношение част – цяло или принадлежност) също биха били сравнително малко.

Като глаголи за притежание (verb.possession) са класифицирани 858 глаголни синсета – глаголът за изразяване на притежание *имам* е част от редица синонимни множества, сред които са: синсетът {имам; притежавам; съм с; съм със} с дефиниция ‘проявявам определени качества, характеристики и т.н.’, който отразява орнативното и комитативното значение на притежанието; синсетът {имам; притежавам} с дефиниция ‘представявам собственик, притежател, източник на нещо конкретно или абстрактно, характеризирам се със или съм носител на определено качество, свойство, признак и т.н.’, който отразява общото състояние на притежание (отчуждаемо и неотчуждаемо); и синсетът {притежавам; имам; владея} с дефиниция ‘собственик съм на вещ, имот и т.н.’, който отразява състоянието на собственост, владеене на (най-често отчуждаемо) притежание.

Общото значение {имам; притежавам} е моделирано като каузирано (свързано е с релацията **причинен от**) от понятието, изразявано от синсета {давам; дам} с дефиниция ‘правя така, че някой да получи нещо в конкретен или абстрактен (чувства и т.н.) смисъл’, което пък е хипоним на {прехвърлям; прехвърля} с дефиниция ‘правя нещо да смени собственика си, предавам на друг отговорността за нещо или някого’ (1.).

1. {давам; дам} (verb.possession)

релация **причинява**: {имам; притежавам} (verb.possession)

релация **хипероним**: {прехвърлям; прехвърля} (verb.possession)

Както вече беше посочено, функцията притежател може да се изпълнява от различни обекти – както одушевени, така и неодушевени, като в Уърднет експлицитна референция към притежателя е кодирана в две синонимни множества за съществителни – {собственик; притежател} и {собственик; притежател; владелец; стопанин}, всяко от които е свързано с експлицитна деривационна и агентивна релация с глаголното синонимно множество {притежавам; имам; владея} (2.а. и 2.б.).

2.а. {собственик; притежател} ‘лице, което притежава определена собственост, представляваща източник на доходи, или пълния дял от бизнеса, който развива’ (noun.person)

релация **дериватив**, релация **е агент на**: {притежавам; имам; владея} ‘собственик съм на вещ, имот и т.н.’ (verb.possession)

2.б. {собственик; притежател; владелец; стопанин} ‘човек, който притежава нещо’ (noun.person)

релация **дериватив**, релация **е агент на**: {притежавам; имам; владея} ‘собственик съм на вещ, имот и т.н.’ (verb.possession)

И двата синсета със съществителни за притежатели (заедно с хипонимите си) се отнасят до подсферата на комерсиално притежание на движимо (отчуждаемо) имущество и са класифицирани като човешки същества (чрез примитива noun.person). Освен това глаголният синсет {имам; притежавам} е свързан с агентивна релация със синсета {богаташ; богат човек; заможен човек; състоятелен човек} ‘човек, който притежава значителни финансови средства и/ или друго материално имущество’, който отново е категоризиран като човек (noun.person).

Тези синонимни множества обаче не реферират пряко към притежателя на дадено имущество, а по-скоро кодират функцията притежаващ. Освен това не реферират и към притежателя на характеристика – както посочихме по-горе, такава функция могат да изпълняват множество обекти.

С агентивна релация е дефинирано и отношението между глаголните синсети {давам; дам} ‘правя така, че някой да получи нещо в конкретен или абстрактен (чувства и т.н.) смисъл’ и {подарявам; подаря; давам; дам} ‘поднасям нещо на някого в знак на добри чувства, благодарност или поради други (религиозни, ритуални и др.) подбуди’, от една страна, и синсета {дарител; донор} ‘благодетел, който безвъзмездно предоставя, извършва дарение на собственост или парични средства за определена обществена, политическа, културна кауза, начинание, инициатива’.

Агентивна релация свързва и глаголния синсет {прехвърлям; прехвърля} ‘правя нещо да смени собственика си, предавам на друг отговорността за нещо или някого’ и синсетите за съществителни {прехвърлящ} ‘лице, което преотстъпва свое право, титла, собственост и др. на друго лице или лица’ и {прехвърлител} ‘физическо или юридическо лице, което прехвърля право

на собственост върху имот, титла и пр. на друго лице или лица'. Синсетът {продавач} 'търговец, често и производител, който предлага стоки и услуги директно на потребителите' е свързан с агентивна релация с глаголния синсет {продавам; продам} 'предлагам стоки или услуги срещу определено заплащане', който, от своя страна, също е част от хиперонимо-хипонимното дърво на {прехвърлям; прехвърля}.

Понятието за {придобиващ} 'човек, който получава, най-често постоянно, право на собственост, ползване и др. върху нещо' е свързано с агентивна релация с глаголния синсет {получавам; получа; придобивам; придобия; сдобивам се; сдобия се; добивам; добия} 'влизам във владение на нещо, ставам притежател на нещо в конкретен или абстрактен смисъл или започвам да имам ново качество, свойство и т.н.' и има хипоним {получател; приемател; реципиент} 'лице, което приема или получава нещо', който е свързан с агентивна релация с глаголния синсет {получавам; получа; имам} 'придобивам (или в по-широк смисъл приемам) нещо, което ми се дава, изпраща, осигурява и т.н.'. Така и синсетът {купувач} 'лице, което придобива правото на собственост върху стока или върху извършването на услуга срещу регламентирана с ценоразпис сума' е свързан с агентивна релация с {купувам; купя; закупувам; закупя} 'придобивам стока или ползвам услуга срещу определено заплащане', който е хипоним на {получавам; получа; придобивам; придобия; сдобивам се; сдобия се; добивам; добия}.

Глаголните синсети {продавам; продам} 'предлагам стоки или услуги срещу определено заплащане' и {купувам; купя; закупувам; закупя} 'придобивам стока или ползвам услуга срещу определено заплащане' са свързани помежду си с антонимна релация. Не така стои обаче въпросът със свързаните с тях агентивни съществителни {продавач} и {купувач}, които не са антонимни и по-скоро реферират към функция или роля (а не към човека, който може да я изпълнява).

Всички тези свързани понятия за лица, участващи в ситуацията по трансфер на притежание, са хипоними на хиперонима за понятието {човек; личност; индивид; лице; човешко същество} с дефиниция 'представител на едноименния и единствен вид на същото семейство висши бозайници (Hominidae), различаващ се от останалите животни по силно развития мозък, съзнание, абстрактно мислене, членоразделна реч; движи се с изправено тяло, произвежда оръдия на труда и други артефакти', което наследява характеристики от две други понятия, негови хипероними – {организъм; създание} с дефиниция 'живо същество с единно организирани функции и устройство, разглеждано от гледна точка на биологичната му същност' и {причинител; причина; причинна сила} с дефиниция 'обект, предмет, същество, което предизвиква определено събитие, явление, ефект, въздействие и др.', което отразява свойството им за упражняване на воля и/или контрол. Както ще видим по-нататък обаче, в езика в тази функция се реализират и други понятия, които могат да се приемат като притежаващи способност за

причиняване на определени действия или събития – това са понятията, класифицирани със семантичния примитив *noun.group*, които най-често са част от хиперонимо-хипонимното дърво на синсета {група; групировка} с дефиниция ‘съвкупност от обекти (неодушевени или одушевени), разглеждана в нейната цялост’. Хипоними на {група; групировка} са понятията за групи от хора и за други видове социални групи (като организация, институция, компания и под.).

3. Изразяване на притежание във Фреймнет

Фреймнет (Baker et al. 1998; Fillmore 1976, 1982) е семантико-синтактичен ресурс, в който лексикалните единици са обединени посредством схематични описания, познати като фреймове. Фреймовете представят концептуалната структура на ситуациите и включват участниците в ситуацията, обстоятелствата и други фреймови елементи (Ruppenhofer et al. 2016; за Българския фреймнет (Коева и др. / Коева et al. 2008; Коева / Коева 2010)². Фреймовите елементи могат да са както ядрени (*core*), така и периферни (*non-core*) участници в лексикално-семантичното представяне на ситуацията.

Притежателят във Фреймнет се кодира експлицитно чрез два отделни фреймови елемента – **Притежател (Possessor)** и **Собственик (Owner)**. Фреймовият елемент **Притежател** е стативен участник и е ядрен за семантични фреймове, които кодират притежание на част от цяло. Това са фреймове за съществителни имена, които изразяват притежание на част от тялото или наличие на (неотчуждаема телесна) характеристика – **Части на тялото (Body parts)**, **Белег на тялото (Body mark)**, **Изражение на лицето (Facial expression)**, **Форма на косата (Hair configuration)**, или притежание на неодушевени същности, за които се предполага също задължително присъствие на притежател – като **Инфраструктура (Infrastructure)**, **Войска (Military)**, **Пари (Money)**, **Склад (Store)**. Фреймовият елемент **Принадлежност към място (Container possessor)** участва в изразяването на отношението съдържател – съдържимо. Фреймовият елемент **Украсен човек (Decorated individual)** кодира орнативното отношение между човек и негова характеристика и участва в семантичния фрейм **Декорация на тялото (Body decoration)**.

От своя страна, роднинските взаимоотношения са кодирани чрез семантичния фрейм **Роднинство (Kinship)**, в който участват ядрените фреймови елементи **Ти (Alter)**, **Аз (Ego)** и **Роднини (Relatives)**, като последният съ-

² Българският фреймнет (Коева / Коева 2010) включва и допълнителна информация, като за всеки предикат се определят категориите лице, преходност и вид, плюс синтактична информация за задължителността на експлицитната реализация. Включени са и семантични характеристики (абстрактно, конкретно, одушевено, неодушевено, лице, име и т.н.); квантитативни ограничения (броимо, неброимо и т.н.) и семантичните класове на възможните фреймови елементи.

държа първите два и съответно използването му изключва използването на другите два. Другите лични взаимоотношения (тук са понятия като приятел, приятелка, годеник, годеница, овдовял, разведен и под.) са изразени чрез семантичния фрейм **Лични връзки (Personal relationship)** с ядрени фреймови елементи **Партньор1 (Partner1)**, **Партньор2 (Partner2)**, **Партньори (Partners)**, където отново наличието на третия изключва реализацията на първите два.

Фреймовият елемент **Собственик** се свързва с активно, съзнателно притежание и може да е участник в ситуации на трансфер на притежаваното. Той е част от семантичните фреймове **Загуба (Losing)** (и от свързаните с него семантични фреймове **Преди загуба на собственост (Pre lose possession)** и **След загуба на собственост (Post lose possession)**). Още три фреймови елемента изразяват участник в ситуация на придобиване или загуба на власт/права върху притежаваното – това са фреймовите елементи **Заемодател (Lender)**, **Първоначален собственик (Original owner)** и **Собственик на бизнес (Proprietor)**.

В останалите семантични фреймове, свързани по различен начин с изразяването на притежание (което фигурира по един или друг начин и в дефинициите им), участват и редица други фреймови елементи, които могат да бъдат интерпретирани като притежатели при описанието на ситуацията.

Фреймовият елемент **Носещ (Wearer)** се изразява от одушевен обект (със семантичен тип Съзнателен (Sentient), който носи дреха и който е притежател на **Част на тялото (Body part)**, като и двата фреймови елемента са ядрени за семантичния фрейм **Носене (Wearing)**.

Фреймовият елемент **Пациент (Patient)** е ядрен за фрейма **Причина за промяна на сила (Cause change of strength)**, където може да се интерпретира като притежателя на силата, който претърпява промяна или е подложен на нея. Подобна интерпретация е в сила и за фрейма **Лекуване (Cure)**, където подлежащият на лечение може да се интерпретира като притежател на (или засегнат от) болестта.

Фреймовият елемент **Реципиент (Recipient)** е ядрен за група фреймове, които описват действия, свързани с трансфер на притежание, като **Прехвърляне (Transfer)**, с поддействия като **Изпращане (Sending)**, **Подаване (Passing)**, **Доставяне (Delivery)** и **Снабдяване (Supply)**; **Даване (Giving)**, **Получаване (Receiving)**, **Вземане (Getting)**, но и **Нагрупване (Amassing)**, **Загуба на притежание (Lose possession)**, **Стиснатост (Stinginess)**, **Отказ от притежание (Surrendering possession)**. Като неядрен елемент **Реципиент** участва в семантичните фреймове за **Ситуация преди даване (Pre giving)** и **Ситуация преди прехвърляне (Pre transfer)**, както и в **Изпратени обекти (Sent items)**. В нито един от тези фреймове обаче елементът **Реципиент** не е ограничен като характеристики (тоест не е нито Съзнателен, нито Човек).

Купувач (Buyer) и **Продавач (Seller)** са фреймови елементи на семантични фреймове, които описват комерсиално отношение на трансфер като (търговско) купуване (**Търговия – купуване (Commerce buy)**) и **Търговия –**

продаване (Commerce sell). И двата семантични фрейма имат ядрен фреймов елемент **Купувач**, като обаче само вторият има и **Продавач**. Във фрейма **Търговия – купуване** може да се реализира неядрен фреймов елемент **Реципиент**, който реферира към потенциален получател, различен от **Купувач** – тоест **Купувач** може да се възприема като временен притежател, купувач обекта в полза на (потенциален) друг притежател. Фреймовият елемент **Вносител (Importer)** е елемент от фрейма **Внасяне (Importing)**, а **Потенциален реципиент (Potential recipient)** е елемент от фрейма **Предлагане (Offering)** – доколкото обектът може да бъде предлаган на някого, но не е задължително той да го получи. Фреймовият елемент **Купувач (Purchaser)** е ядрен за семантичния фрейм **Пазаруване (Shopping)** и може да се отнася за временен притежател (както ще видим и по-нататък).

Ситуацията на трансфер на притежанието се кодира от дитранзитивни предикати. При тях обектът се трансферира, прехвърля или премества от един обект във или към друг обект – така наречения от Ст. Пинкър „нов“ притежател (Pinker 1989: 48), който според изследователите се моделира и като локация на притежанието. Тоест в тази ситуация участват двама „притежатели“ или два обекта, които биха могли да се интерпретират като локация на притежанието на трансферирания обект. Тук трябва да се направи уточнението, че реципиентът (получателят) не е задължително да е притежател, поне не и постоянен – например и купувачът, и продавачът могат да са временни притежатели на обекта, така че фреймовите елементи, изразяващи реципиентна функция, не са свързани с фреймовите елементи **Притежател** или **Собственик**.

Следващата част представя някои наблюдения върху реализациите именно на фреймовите елементи **Реципиент**, **Купувач** и **Продавач** при предикатите, изразяващи комерсиален трансфер на притежание (които са разглеждани у (Коева / Коева 2010: 42 – 45).

4. Временни комерсиални притежатели

За да се наблюдава реализацията на съответните значения на глаголите за трансфер на притежание, в които се реализира фреймовият елемент **Реципиент**, са извлечени ограничен брой корпусни примери (до 1000) чрез приложението СкечЕнджин (SketchEngine) (Kilgariff et al. 2014). Примерите са ръчно филтрирани чрез отстраняване на значения, различни от разглежданите (например *доставям* в съчетание със съществителни за емоция като *радост*, *удоволствие*, *наслада*, *наслаждение*, както и със съществителни като *информация*, *сведения*; глаголите *купувам*, *продавам*, *пазарувам* в съчетание с *гласове*, *избиратели* и под.), като са избрани примери, илюстриращи общото значение ‘осъществявам трансфер на (материален) обект с търговски или други комерсиални цели’.

В наблюдаваните примери много често (всъщност много по-често при предикатите за комерсиален трансфер на притежание) експлицитно се ре-

лизират, освен ядрените фреймови елементи като **Стока (Goods)** или **Тема (Theme)**, предимно периферни фреймови елементи като **Място (Place)**, **Предназначение (Purpose)**, **Цел (Goal)**, **Време (Time)**, **Пари (Money)** и др.

4.1. Реципиент

Семантичният фрейм **Доставяне** има четири ядрени фреймови елемента – **Доставчик (Deliverer)**, **Цел, Реципиент** (като при наличие на **Реципиент** се изключва **Цел**) и **Тема**, и седемнадесет периферни фреймови елемента, сред които е и **Бенефициент** с дефиниция „Лицето, в чиято полза се осъществява доставката“.

В извадка от 471 примера от корпуса в 200 е експлицитно реализиран фреймовият елемент **Реципиент**, който се изразява от: местоименна клитика (включително възвратната *си*) – в 137 примера (3.а.); предложни фрази, въведени най-често с предлога *на* – в 63 примера (3.б.); много по-рядко с предлога *за*, за който се предполага, че въвежда **Бенефициент** – едва в 7 примера (3.в.).

- 3.а. ... [той]_{Доставчик} [ни]³_{Реципиент} го ДОСТАВЯШЕ право на [нас]_{Реципиент} ...
 3.б. Дали [Коуди]_{Доставчик} не беше ДОСТАВЯЛ [хероин]_{Тема} [на тоя човек]_{Реципиент} ?
 3.в. ... когато [Патрик]_{Доставчик} ДОСТАВЯЛ [един отличен жребец]_{Тема} от родната си Ирландия [за дъщерята на едрия земевладелец]_{Бенефициент}.

Сред изразяваните периферни фреймови елементи са **Място** (4.а.), **Предназначение** (4.б.), **Източник** (4.в.), който може да бъде и „стар“ притежател, и други.

- 4.а. А има само [двама пощальони]_{Доставчик}, които са ДОСТАВЯЛИ [писмата]_{Тема} [на Уест стрийт]_{Място} [през декември]_{Време} ...
 4.б. ... [хора]_{Доставчик}, които ДОСТАВЯХА [добитък]_{Тема} [за пировете ви]_{Цел} ...
 4.в. [Всичко необходимо за живота на хората там]_{Тема} се ДОСТАВЯШЕ [от другите пет континента]_{Източник}.

4.2. Купувач

Семантичният фрейм **Търговия – купуване** има два ядрени фреймови елемента – **Купувач** и **Стока**, и тринадесет периферни, сред които фигурира и **Реципиент** (в този фрейм няма фреймов елемент **Бенефициент**) с дефиниция „Лицето, за което купувачът купува стоката“.

³ В примерите са отбелязани ядрените фреймови елементи, а от периферните – тези, които се разглеждат в съответната част на текста.

В извадка от 970 примера само в 73 допълнително е акцентирано върху **Купувача** (в 72 референтът е или човек, или група от хора), като най-често чрез възвратна клитика (*купувам си X*) се акцентира върху самия **Купувач** (обикновено в позицията на субект) като „нов“ притежател на обекта – в 51 примера (5.а.). В останалите случаи се среща (периферен фреймов елемент) **Реципиент** (в бенефициентна функция, т.е. някого, в чиято полза се купуват стоките), различен от **Купувача** – 17 примера с местоименна клитика, различна от възвратната (5.б.), както и 4 предложни фрази, въведени с предлога *на* (5.в.), и 4 предложни фрази, въведени с предлога *за*, включително и с референт, различен от човешко същество, но все пак отнасящ се до търговско юридическо лице (5.г. и 5.д.).

- 5.а. *Те [си]_{Купувач} КУПУВАТ [дрехи]_{Стока} от сергиите ...*
 5.б. *Вернер винаги ми донасяше кафе в междучасията, КУПУВА-ШЕ [ми]_{Реципиент} [виенски кифлички с шоколад]_{Стока} ...*
 5.в. *[Майката]_{Купувач} КУПУВА [на малкото си болно момченце]_{Реципиент} [кутийка боички]_{Стока} ...*
 5.г. *За целта [мъжете]_{Купувач} КУПУВАТ [цветя]_{Стока} [за съпругите и колежките си]_{Реципиент} ...*
 5.д. *Аничке, виж таз сметка, дето съм КУПУВАЛ [едно-друго]_{Стока} [за банката]_{Реципиент} ...*

В извлечените примери много често се реализират другите периферни фреймови елементи като **Място** (6.а. и 6.б.), **Предназначение** (6.б.), **Пари** (6.в.), **Източник** (който може да е и „стар“ притежател и **Продавач**) (6.г.).

- 6.а. *Освен, че ги КУПУВАХА [в безистена и магазините по Широк сокак]_{Място} ...*
 6.б. *Те били КУПУВАНИ [за разни стопански и битови нужди] ^{Предназначение} както [на вътрешния пазар]_{Място}, така и [по международните робски тържища из Леванта]_{Място}.*
 6.в. *... КУПУВА [остарели котки]_{Стока} [по четвърт долар]_{Пари} ...*
 6.г. *Три години съм КУПУВАЛ [от баща му]_{Източник} [вино]_{Стока} ...*

4.3. Продавач

Семантичният фрейм **Търговия – продаване** има три ядрени фреймови елемента – **Купувач**, **Стока** и **Продавач** – и четиринадесет периферни.

В извадка от 895 корпусни примера само в 115 има експлицитно изразен реципиент (**Купувач**, който може да има и бенефициентна интерпретация (8.б.), който може да е както човешко същество (7.а.) – в 14 примера, така и по-често – група от лица (групи от хора – в 65 примера, или други организации – в 36 примера) (7.б. и 7.в.).

- 7.а. ... [човек]_{Продавач} ПРОДАВА [своя бѣбрек]_{Стока} [на богат рецепиент]_{Купувач} поради бедността си ...
- 7.б. ... и започнал да посещава магазините в Бейрут, където ПРОДАВАЛИ [монети]_{Стока} [за колекционерите]_{Купувач} ...
- 7.в. ... после ПРОДАВА [продуктите от тези изследвания]_{Стока} [на слаборазвитите страни]_{Купувач} ...

При тези предикати много по-често се реализират периферните фреймови елементи, сред които са **Място** (8.а.), **Пари** (8.б.), **Предназначение** (8.в.), **Начин** (8.г.).

- 8.а. [В този град]_{Място} те ПРОДАВАЛИ [селскостопанските си произведения]_{Стока} и оттам купували захар, сол, газ, покъщнина.
- 8.б. От казаните се получавали [600 оки масло]_{Стока}, което ПРОДАВАЛИ [за 600 златни лири]_{Пари}.
- 8.в. ... но веднъж той отишъл да ПРОДАВА [жито]_{Стока} [за данък]_{Предназначение} ...
- 8.г. ... [съсед]_{Продавач}, който ПРОДАВА [ориза]_{Стока} [на загуба]_{Начин} и все пак печели, помагайки си с теглилката.

4.4. Купувач

И накрая ще бъде разгледан семантичният фрейм **Пазаруване (Shopping)** с два ядрени фреймови елемента – **Стока** и **Купувач (Purchaser)**, със семантичния тип Съзнателен (в разгледаните по-горе семантични фреймове няма ограничения за характера на **Купувача** или **Продавача**). Сред десетте периферни фреймови елемента на този семантичен фрейм не фигурира нито **Бенефициент**, нито **Реципиент** (различен от **Купувач**).

В извадка от 961 примера в 44 е изразен бенефициент, като в 24 примера върху него се акцентира с възвратна клитика (*пазарувам си*) или е въведен с възвратно местоимение с предлога *за* (*пазарувам за себе си*). В останалите случаи бенефициентът е въведен от лична местоименна клитика, различна от възвратната (4 примера), и по-често от предлога *за* (в 11 примера), отколкото от предлога *на* (само в 2 примера). В повечето примери референтът на този фреймов елемент е човешко същество (9.а. и 9.б.) или група (9.в. и 9.г.).

- 9.а. Новият генерален директор ли го взе за себе си да му тупа килимите вкъщи и да [му]_{Бенефициент} ПАЗАРУВА, друг някой ли го сви (...)
- 9.б. Днес обаче ПАЗАРУВАМЕ [за Люк]_{Бенефициент}, а не [за мен]_{Бенефициент}.
- 9.в. (...) когато брат ми беше в града, гуляеше или ходеше да ПАЗАРУВА [за нас]_{Бенефициент} [разни неща]_{Стока}.
- 9.г. (...) а защо не например да се ПАЗАРУВА [на възрастни родители]_{Бенефициент} и др.

В тези примери отново много по-често се реализират периферни фреймови елементи, сред които **Място** (10.а.), включително **Източник** (10.б.), **Предназначение** (10.в.), **Начин** (10.г.).

10.а. ... и ПАЗАРУВАШЕ [хранителни стоки]_{Стока} [в единствения супермаркет]_{Място}.

10.б. ... [Мати Девор]_{Купувач} е принудена да ПАЗАРУВА [преоценени дрехи]_{Стока} [от „Кеймарт“]_{Източник} ...

10.в. Отсега ПАЗАРУВА [за Коледа]_{Предназначение}.

10.г. [Майка ѝ]_{Купувач} ПАЗАРУВАШЕ [на едро]_{Начин}, въобразявайки си, че пести пари.

В повечето примери обаче няма експлицитно изразен обект на действието **Стока** (това е фреймов елемент, който се интерпретира и като изразен по подразбиране от лексикалното значение на глагола, а оттам и незадължителен за експлициране), за разлика от предикатите, които реализират останалите три разгледани фрейма. Освен това в реализациите наблюдаваме и други (периферни) фреймови елементи, които не са дефинирани като такива – като например **Бенефициент** при семантичния фрейм **Пазаруване**.

От извлечените реализации на фреймовите елементи в разгледаните фреймове може да се предположи, че съществителните, удовлетворяващи изискванията на съответните глаголи (за (комерсиален) трансфер на притежание) за обекти, между които се извършва трансферът, реферират към широк кръг от понятия, които могат най-общо да бъдат определени като лица – физически или юридически. Последните могат да бъдат представлявани от физическо лице, което може да проявява воля и да има контрол върху обекта, влизащ в ролята на (временно или постоянно) притежание. От друга страна, както вече беше отбелязано в предходните части, в отношение на стативно притежание или принадлежност е възможно да влизат дори и същности, които не могат да проявяват воля (като артефакти, растения, абстрактни характеристики и под.).

5. Заключение

Доколкото функцията на т.нар. притежатели не може да се положи в ясно дефинируеми граници, е трудно да се определят класовете съществителни имена, рефериращи към възможни същности, които могат да изпълняват тази функция. Фокусът върху участниците в ситуацията по трансфер на притежание, поставен в настоящата студия, е свързан с разбирането, че обектите от сферата на отчуждаемата посесивност, които могат да подлежат на трансфер, са участници в отношение на притежание – тоест именно в тези ситуации очакваме реализация на притежатели, които могат по собствена воля да се сдобият с нещо (да станат притежатели) и да се откажат от / да се разделят с нещо (да спрат да са притежатели).

Във всички тези случаи в корпусните примери, които бяха разгледани, се реализират съществителни имена, свързани с две понятия в Уърднет. Първото понятие {човек; личност; индивид; лице; човешко същество} е част от хиперонимо-хипонимното дърво на {физическа същност} ‘същност, която има физическа форма, материално проявление или чието съществуване е наблюдаемо във физическия свят’. Второто понятие {група; групировка} ‘съвкупност от обекти (неодушевени или одушевени), разглеждана в нейната цялост’ обаче е хипоним на {абстракция; абстрактна представа; абстрактно понятие} ‘форма на познание, основана на съществените свойства и отношения на предмета и получаване на общо понятие за него’. За да бъдат свързани тези две понятия, трябва допълнително да бъдат дефинирани характеристики, които да определят способността им да изпълняват действия, изискващи упражняване на воля и на контрол. В разглежданите примери не се разграничава и значението на семантичната характеристика постоянен или близък контакт при отчуждаемите притежания.

Благодарности: Изследването е извършено в рамките на проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания“ (Договор № КП-06-Н 50/1 от 30.11.2020 г.).

Цитирана литература / References

- Коева и др. 2008: Коева, С., Р. Влахова, Р. Декова, П. Несторова, А. Атанасов. *Българският ФреймНет: семантико-синтактичен речник на българския език*. Съст. С. Коева. София: Институт за български език „Проф. Любомир Андрейчин“. (Koeva et al. 2008: Koeva, S., R. Vlahova, R. Dekova, P. Nestorova, A. Atanasov. *Balgarskiyat FreymNet: semantiko-sintaktichen rechnik na balgarskia ezik*. Sast. S. Koeva. Sofia: Institut za balgarski ezik “Prof. Lyubomir Andreychin”.)
- Коева 2010: Коева, С. *Българският ФреймНет*. София: Институт за български език „Проф. Любомир Андрейчин“. (Koeva 2010: Koeva, S. *Balgarskiyat FreymNet*. Sofia: Institut za balgarski ezik “Prof. Lyubomir Andreychin”.)
- Лесева, Стоянова 2020: Лесева, С., И. Стоянова. Отвѣд лексикалните ресурси: валидиране на концептуалното описание в корпусни данни. – В: Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“ (София, 2020 г.). Т. 2. София: Издателство на БАН „Проф. „Марин Дринов“, с. 241 – 249. (Leseva, Stoyanova 2020: Leseva, S., I. Stoyanova. *Otvad leksikalnite resursi: validirane na kontseptualnoto opisanie v korpusni dannii*. – In: *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik “Prof. Lyubomir Andreychin”* (Sofia, 2020 g.). Т. 2. Sofia: Izdatelstvo na BAN “Prof. „Marin Drinov”, pp. 241 – 249.)
- Baker et al. 1998: Baker, C., C. Fillmore, J. Lowe. The Berkeley FrameNet project. – In: *COLING 1998, Vol. 1: The 17th International Conference on Computational Linguistics*. Montreal: Association for Computational Linguistics.

- Benveniste 1966: Benveniste, É. *Problèmes de linguistique générale*. Paris: Gallimard.
- Fellbaum et al. 2009: Fellbaum, C., A. Osherson, P.E. Clark. Putting Semantics into WordNets Morphosemantic Links. – In: *Human Language Technology. Challenges of the Information Society. LTC 2007. Lecture Notes in Computer Science*, vol. 5603. Springer, Berlin, Heidelberg, pp. 350 – 358.
- Fillmore 1976: Fillmore, C. J. Frame Semantics and the Nature of Language. – *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, vol. 280, no. 1, pp. 20 – 32.
- Fillmore 1982: Fillmore, C.J. Frame Semantics. – In: *Linguistics in the Morning Calm*. The Linguistic Society of Korea, Seoul: Hanshin, pp. 111 – 137.
- Heine 1997: Heine, B. *Possession: Cognitive Sources, Forces, and Grammaticalization*. Cambridge: Cambridge University Press.
- Herslund, Baron 2001: Herslund, M., I. Baron. Introduction: Dimensions of Possession. – In: M. Herslund, I. Baron, F. Sørensen (Eds.). *Dimensions of Possession*. Amsterdam: John Benjamins, pp. 1 – 25.
- Kilgarriff et al. 2014: Kilgarriff, A., V. Baisa, J. Bušta, M. Jakubíček, V. Kovář, J. Michelfeit, P. Rychlý, V. Suchomel. The Sketch Engine: Ten Years on. – *Lexicography*, vol. 1, no. 1, pp. 7 – 36.
- Koeva 2021a: Koeva S. The Bulgarian WordNet: Structure and Specific Features. – *Papers of BAS. Humanities and Social Sciences*, vol. 8, no. 1, pp. 47 – 70.
- Koeva 2021b: Koeva, S. Towards Expanding WordNet with Conceptual Frames. – In: *Proceedings of the 11th Global WordNet Conference*. University of South Africa (UNISA): Global WordNet Association, pp. 182 – 191.
- Levin 1993: Levin, B. *English Verb Classes and Alternations*. University of Chicago Press: Chicago.
- Miller et al. 1993: Miller, G., R. Beckwith, C. Fellbaum, D. Gross, K. Miller. *Introduction to WordNet: an On-line Lexical Database. Five Papers on WordNet*. Princeton, NJ: Princeton University.
- Payne, Barshi 1999: Payne, D., I. Barshi (Eds.). *External Possession*. Amsterdam: John Benjamins.
- Pinker 1989: Pinker, S. *Learnability and Cognition. The Acquisition of Argument Structure*. Cambridge, MA: MIT Press.
- Rappaport Hovav, Levin 2008: Rappaport Hovav, M., B. Levin. The English Dative Alternation: The Case for Verb Sensitivity. – *Journal of Linguistics*, vol. 44, no. 1, pp. 129 – 167.
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016). < <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf> > [03.03.2024]
- Seiler 1983: Seiler, H.-J. 1983. *Possession as an Operational Dimension of Language*. Tübingen: Gunter Narr Verlag.
- Stassen 2009: Stassen, L. *Predicative Possession*. Oxford: Oxford University Press.

ON THE NATURE OF POSSESSORS

Tsvetana Dimitrova

Institute for Bulgarian Language “Prof. Lyubomir Andreychin”

cvetana@dcl.bas.bg

Abstract. The article reviews the expression of possessors with nouns referring to entities that may satisfy the requirements of frame elements which are part of the semantic frames in FrameNet. The frame elements realised with the main verbs for transfer of possession in a limited number of corpus examples are considered. Apart from the frame element (referring to the object of the transfer) and the two entities between which/whom the transfer of possession takes place, non-core frame elements are also explicitly realised. The sets of appropriate nouns and their features are defined based on the realisation of the frame elements in the corpus examples.

Keywords: *possession, WordNet, FrameNet*

Tsvetana Dimitrova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52 Shipchenski prohod Blvd., Bl. 17
Sofia 1113, Bulgaria
<https://orcid.org/0000-0002-8972-435X>

ЗА СЕМАНТИЧНИЯ КЛАС ГЛАГОЛИ ЗА КОНТАКТ ЧРЕЗ ПРЕМЕСТВАНЕ

Мария А. Тодорова

Институт за български език „Проф. Любомир Андрейчин“

maria@dcl.bas.bg

Резюме. Студията предлага описание на група глаголи от Уърднет със семантичния клас (примитив) „глаголи за контакт“, които принадлежат към множеството от общоупотребимата лексика в български. Описанието се фокусира върху глаголите за контакт, комбиниран с преместване, които включват глаголи за прикрепване, отделяне, поставяне, премахване, пълнене и изпразване. Предложено е семантико-синтактично описание, което обхваща фреймовите елементи от съответните семантични фреймове на Фреймнет и съответстващите им синтактични реализации. Разглеждат се ситуации, представени от семантичните фреймове **Прикрепване, Отделяне, Запълване, Изпразване, Манипулация и Засягане**. Анализирания информация се основава на ръчна проверка, редакция и допълване на полуавтоматичното съотнасяне на два големи семантични езикови ресурса – Уърднет и Фреймнет. Предложени са подкласове на глаголите за контакт чрез преместване според локацията на физическия контакт.

Ключови думи: глаголи за контакт и преместване; семантико-синтактично описание

1. Въведение

Студията предлага семантично описание на група глаголи от Уърднет със семантичния клас (примитив) „глаголи за контакт“, които принадлежат към множеството от общоупотребимата лексика в български. Разглежда се общата семантика на класа и подходите за нейното представяне в езиковедските изследвания (Fillmore 1970; Levin 1993; Gao, Cheng 2003 и др.). Глаголите за контакт се разглеждат в два основни класа: глаголи за физически контакт чрез движение и глаголи за физически контакт при състояние, като

* Мария А. Тодорова. За семантичния клас *глаголи за контакт чрез преместване*. – В: Св. Коева (ред.) *Семантични студии. Синтактично и семантично описание на български глаголи*. София: Издателство на БАН „Проф. Марин Дринов“, с. 134 – 153. <https://doi.org/10.7546/SemanticStudies2024.06>

настоящото описание се фокусира върху една от подгрупите на глаголите за контакт чрез движение. Изборът на данните и тяхното описание се основава на предположението, че семантичните фреймове на глаголите за контакт включват действия или събития, при които два или повече обекта влизат във физически контакт един с друг. Коментирани са само най-честите случаи с по-голям обхват в Българския уърднет (Булнет) (Коева 2021a).

Уърднет (Miller 1995; Fellbaum 1999) и Фреймнет (Fillmore 1982; Baker et al. 1998) са големи лексикално-семантични ресурси, които предоставят семантична информация за голямо и разнообразно множество от глаголи (и други части на речта). Уърднет представлява многоезикова мрежа от синонимни множества (синсети), свързани посредством семантични отношения като хиперонимия, антонимия и т. н., и предлага групиране на глаголите и съществителните в семантични класове. Фреймнет представя семантиката на лексикалните единици чрез семантични фреймове, описващи ситуации и техните компоненти (елементи на семантичния фрейм) (Fillmore 2007: 29).

Избраният подход за изследване е предпоставен от идеята на Джакендоф, че смисълът на дадено изречение е изграден от концептуалните структури на лексикалните единици, съставляващи изречението (Jackendoff 1997: 48). Тази гледна точка е свързана с ограниченията върху комбинирането на дадена дума с други думи, наричани *селективни ограничения*, които според Чомски (Chomsky 1965: 95) се дължат на синтактичните свойства на думите. За да групираме избраните глаголи за контакт в подтипове въз основа както на реализацията на техните фреймови елементи, така и на тяхната лексикална семантика, приемаме, че глаголите с подобни значения споделят общи валентни модели (valence patterns) и общи модели на реализация на фреймовите елементи на синтактично равнище.

2. Глаголи за контакт

2.1. Обща семантика на класа

Глаголите за контакт принадлежат към множество от глаголни предикати, за които е характерен концептуалният ядрен признак **ФИЗИЧЕСКИ КОНТАКТ** (Juffs 1996: 47). В Уърднет те са включени в един от 15-те лексикографски файла, в които глаголите са групирани според семантичната област, към която се отнасят, и са определени като „глаголи за докосване, удране, връзване, копаене“ (Miller et al. 1990: 47).

Множеството на глаголите за контакт обхваща най-големия семантичен клас глаголи в Уърднет, състои се от повече от 820 синонимни множества, включващи глаголи за действия, състояния и събития, които споделят семантичния компонент **КОНТАКТ (СЪПРИКОСНОВЕНИЕ)**. Този тип глаголни синонимни множества са част от йерархия, изградена чрез хипонимно-хиперонимните (тропонимните) релации, чрез които се въвеждат и

редица различни прояви на общ семантичен признак (Fellbaum, 1999: 94). Семантичната дефиниция на класа е твърде широка и всъщност не обобщава семантиката на всички глаголи, които съдържа. В български, а и в други езици семантичният клас на глаголите за контакт е по-слабо проучен в сравнение с другите класове глаголи. Глаголи за контакт са обект на изследване за английски (Fillmore 1970, Levin 1993, Fellbaum 1990), китайски (Gao, Cheng 2003), шведски (Viberg 2004).

Описанието на класа се основава на първоначално разделение, основано на йерархичната организация на семантичните фреймове (Ruppenhofer et al. 2016), на групирането на глаголите в аспектуални класове: **ДЕЙСТВИЯ (ACTIVITIES)** и **СЪСТОЯНИЯ (STATES)**, предложено от Вендлер (Vendler 1967: 100), както и на идеята за допълващи се понятия или семантични атрибути в организацията на значението (Löbner 2011: 2). Различните глаголи за контакт кодират повече от един семантичен атрибут и в зависимост от фокуса на класификацията могат да бъдат причислени към повече от един семантичен клас. Това може да се види при изследване на семантичните фреймове на класовете глаголи за промяна, глаголи за движение и глаголи за комуникация в Уърднет.

Общата ситуация, представяна от глаголите за контакт, включва действия или състояния, при които две или повече същности влизат в съприкосновение една с друга. Тази ситуация отговаря на концептуален фрейм – „абстрактна структура, която определя семантичната и синтактичната съвместимост между глаголни предикати и техните аргументи, съществителни имена. Конкретна концептуална рамка е свързана със семантичен клас, който изразява неговите общи семантични свойства, представени от набор от глаголи, организирани в синонимните множества на Уърднет и описани от набор от фреймови елементи“ (Коева 2021b: 182). Разглежданите концептуални фреймове, в които участват избраните глаголи представят различни прояви на физически контакт между обекти или субекти и обекти. Основните участници в ситуацията са **СЪЩНОСТИ В КОНТАКТ**² или в съприкосновение, представени от ядрени или периферни фреймови елементи. Ядрените участници в ситуацията са поне две същности, често наричани субект (този, който инициира съприкосновението) и обект (този, с който се осъществява съприкосновението). Някои глаголи за контакт уточняват посоката на контакта и естеството на контакта, което може да варира в широки граници: чрез периферните фреймови елементи **НАСОЧЕНОСТ** и **НАЧИН НА КОНТАКТ**.

Много от глаголите за контакт са описани със семантични фреймове от Фреймнет, кодиращи промяна, движение, състояние и комуникация – **Движение (Motion)**, **Предизвикване на промяна (Cause change)**, **Прегърпяване на промяна (Undergo change)**, **Състояние (State)**, **Комуникативен процес**

² Фреймовите елементи във Фреймнет биват ядрени – съществени за реализирането на всеки уникален семантичен фрейм и периферни – свързани с допълнителни описания и обстоятелства (време, място и т. н.) (Fillmore 2007: 133).

(Commutative process), Некомуникативен процес (Noncommutative process). Те са изключени от обхвата на настоящата студия, която разглежда само глаголи в пряко значение, назоваващи действия, в които се включва физическо съприкосновение на участниците в ситуацията.

Приема се разделението на предикатите за контакт в две подгрупи, които комбинират семантичните компоненти ФИЗИЧЕСКИ КОНТАКТ и ДВИЖЕНИЕ и ФИЗИЧЕСКИ КОНТАКТ и СЪСТОЯНИЕ. Най-типични за семантиката на класа фреймови елементи в избраните семантични фреймове от Фреймнет са **Обект (Item)**, **Тема (Theme)**, **Източник (Source)**, **Част на тялото (Body part)**, **Цел (Goal)** и **Инструмент (Instrument)**.

2.2. Класификации

Групирането на глаголите в семантични класове варира в зависимост от възприетата теория и фокуса на описанието. Изследвайки моделите на редуване на английски, Ч. Филмор (Fillmore 1972: 125) дефинира два големи класа: глаголи за прекъсване (BREAK verbs) и глаголи за удар (HIT verbs), чиито членове споделят общи семантични компоненти и модели на поведение.

В своя анализ на валентността на глаголи от различни езици Б. Левин (Levin 1993) дефинира разграничение на видовете глаголи чрез семантичните компоненти НАЧИН и РЕЗУЛТАТ. Глаголи като *удрям* (hit), които описват осъществяване на повърхностен контакт с обект чрез силен удар, са глаголи, описващи НАЧИН за потенциално увреждане на обекти (Levin 2015: 1634).

Отделни подтипове на класа също са описани от някои автори: глаголи за физически контакт (Gao 2001) и глаголи за удар и въздействие (Gao, Cheng 2003). Б. Левин (Levin 1993: 148 – 156) разделя класа на глаголите за контакт чрез въздействие на: глаголи за удар, глаголи за пляскане и неагентивни глаголи. Р. Декова (Dekova 2006, 2011) изследва 20 глагола за физически контакт в български и английски, разделени в няколко подгрупи. М. Димитрова-Вълчанова и Р. Декова (Димитрова-Вълчанова, Декова / Dimitrova-Valchanova, Dekova 2009) представят корпус и емпирично изведена класификация на глаголите за контакт чрез въздействие, използвайки формализма на Знаковия модел (Hellan, Dimitrova-Valchanova 1995).

Въз основа на хиперонимните връзки в Уърднет Милър и др. (Miller et al. 1990: 59) дефинират следните подгрупи глаголи за контакт:

- а) Глаголи, кодиращи сила, интензивност или повторение на действието (*удрям*);
- б) Глаголи за държане (*хващам, стискам, щипя*) и докосване (*пипам, поглаждам*);
- в) Глаголи, включващи инструмент или материален аргумент (*боядисвам*);
- г) Глаголи, включващи аргумент – част от тялото: рамо (*нарамвам*); пръст (*пипвам*).

Предходната работа върху семантичното описание на българските глаголи включва анализ на глаголи за промяна (Лесева, Стоянова / Leseva, Stoyanova 2021) и глаголи за комуникация (Кукова 2020), глаголи за движение (Костова / Kostova 2010), предикати за психично състояние (Стаменов / Stamenov 2021, Тишева / Tisheva 2021, Джонова / Dzhonova 2008) и описание на синтактични трансформации на български глаголи (Коева / Koeva 2021б, 2021а, 2022). Различни етапи от изследването на семантичните характеристики и селективните ограничения, имащи отношение към семантичното описание на българските глаголи и техните фреймови елементи, са разгледани от Лесева и колектив (Leseva et al. 2019, Leseva et al. 2020).

Доколкото ни е известно, множеството от глаголи, означаващи физически контакт, досега не е описвано като отделен семантичен клас за български език.

3. Глаголи за контакт чрез преместване

Разглежданите глаголи са подклас на избрано множество високочестотни български глаголи, които принадлежат към общата лексика. Подборът им се основава на принадлежността им към глаголни подкласове в Уърднет (Miller 1995) и Фреймнет (Fillmore 1982). В Уърднет³ те са категоризирани като „глаголи за контакт“, а във Фреймнет са представени чрез семантичните фреймове⁴: **Прикрепване (Attaching) и Detaching (Отделяне), Наместване (Adjusting), Запълване (Filling) и Изпразване (Emptying), Манипулация (Manipulating), Засягане (Impact)**.

Въз основа на семантичните фреймове, лексикалните единици, които ги предизвикват, релациите между фреймовете и техните фреймови елементи, както и на селективните особености, свързани със синтактичната реализация на конкретните лексикални единици, са предложени семантични подвидове на разглежданите глаголи. Те принадлежат към класа глаголи, означаващ физическо взаимодействие, което възниква в резултат на движение. Подкласът *глаголи за контакт чрез преместване* се характеризира с типове движение в комбинация с начина на контакт. Този подклас включва глаголи, означаващи прикрепване, отделяне, поставяне, премахване, пълнене и изпразване. Те споделят общи фреймови елементи и ограничения въз основа на мястото на физическия контакт, като включват придвижването на **Обект** към дадена **Цел**. Ядрените фреймови елементи споделят общи абстрактни характеристики – **Агент (Agent)** е волеви участник; **Причина**

³ Данните от Българския уърднет, успореден с уърднетите за английски и други езици чрез междуезиковия индекс ILL, могат да се проследят на <http://dcl.bas.bg/bulnet/>.

⁴ Използвани са данните от Фреймнет, налични към март 2024 на <https://framenet.icsi.berkeley.edu/>.

(Cause) е физическо лице или събитие, **Обект**, **Контейнер (Container)**, **Конектор (Connector)** и **Цел** са физически обекти. Според точката на физически контакт се определят три основни подгрупи: глаголи за контакт върху или по протежение на повърхност, глаголи за контакт с контейнер (вместител), глаголи за контакт с тяло. Разглеждат се варианти на семантичните фреймове **Запълване** и **Изваждане** за две от подгрупите: **Запълване1** и **Изваждане1** при глаголи за контакт върху или по протежение на повърхност и **Запълване2** и **Изваждане2** при глаголи за контакт с контейнер.

3.1. Глаголи за контакт върху или по протежение на повърхност

Глаголите за контакт върху или по протежение на повърхност описват действия, при които **Обект/Тема** влиза в контакт с, движи се по протежение на или следва контура на повърхност. Затова характерна при тези глаголи е повърхността, където възниква контактът или по която се извършва движението, представена от фреймовите елементи **Цел** и **Конектор**. Повърхността е ключов компонент в разбирането на пространствения аспект на действието, описано от тези глаголи. Според начина на контактното движение тези глаголи означават директен физически контакт между два обекта като *докосвам*; лек физически контакт, който включва непряко физическо взаимодействие, като *четкам*, *ожулкам*; нежно съприкосновение като *галя*, *потупвам*, *сгушвам се* или изследващ контакт като *опипвам*, *натискам*.

Към тази група се включват и глаголите, които изразяват физически контакт, осъществен при покриване или отстраняване на **Тема** – някаква субстанция, вещество от дадена повърхност.

По-долу са представени семантичните фреймове, описващи глаголите в рамките на този подклас, техните фреймови елементи, синтактичната им реализация, примери за употреба и селективни специфики.

3.1.1. Семантичен фрейм **Прикрепване**

Дефиниция на семантичния фрейм: „Агент прикрепя **Обект** към друг **Обект** или към **Цел**, като използва **Конектор**“.

Фреймови елементи: **Агент**, **Обект**, **Цел**, **Конектор**

Глаголи: коренът на поддървото в Уърднет⁵ eng-30-01354673-v: {връзвам}, ‘прикрепям, сглобявам или по друг начин осъществявам постоянен контакт между две или повече неща или части’ и хипонимите⁶: {събирам се; събера се; съединявам се; съединя се} ‘свързвам се плътно (и трайно)’; {съединявам; съединя; свързвам; свържа} ‘събирам, свързвам две неща

⁵ Коренът е взел в структурата на Уърднет, представен от синонимно множество, чието значение представя категория, под която са включени по-специфични значения.

⁶ Избрани са само хипоними, които принадлежат към семантичния клас глаголи за контакт в Уърднет и които са част от общоупотребимата лексика.

чрез прилепяне плътно едно до друго така, че да функционират като цяло'; {прикрепвам се; прикрепявам се; прикрепям се; прикрепя се; закрепвам се; закрепям се; закрепя се; залепвам се; залепям се; залепя се; свързвам се; свързва се} 'съединявам се с нещо друго, обикновено с помощта на някакво приспособление, вещество или материал, в резултат от което започвам да се намирам в контакт или образувам физическо и/или функционално цяло с него'; {присаждам; присадя} 'съединявам части от различни растения, като ги накарам да се сраснат, за да подобря характеристиките им и да принеса някакви свойства'; {споявам; споя; припоявам; припоя} 'съединявам части, като използвам разтопена сплав, която след втвърдяването си държи частите заедно'; {заварявам; запоявам; запоя; заваря} 'прикрепвам заедно на две или повече части чрез местно разтопяване и прилагане на натиск'; {прикрепвам; прикрепявам; прикрепям; прикрепя; закрепвам; закрепям; закрепя; закачвам; закача; прикачвам; прикача; закачам} 'съединявам, свързвам две или повече неща'.

Примери⁷:

1.а. [Ани]_{Агент} лесно можеше да **СЪЕДИНИ** [жичките]_{Обект} и да запали мотора.

1.б. [Денис]_{Агент} искаше да **СВЪРЖЕ** [двете тръби]_{Обект} [с дървената гривна]_{Конектор}.

1.в. [Той]_{Агент} **ВЪРЗА** [с връв]_{Конектор} [разваления зъб]_{Обект}.

1.г. [С парче шнур]_{Конектор} [Агент] **ЗАКРЕПИ** [пердетата]_{Обект} [към рамките на прозореца]_{Цел}.

Валентен модел:

| Агент | Обект | Цел | Конектор |
|----------------|---|-----------------------------|--|
| NP (субект) | NP (пряк обект) PP за (непряк обект) | PP към/за (непряк обект) | NP (пряк обект) PP с/със (непряк обект) |

Селективни особености: Общата семантика на фреймовия елемент **Конектор** се изразява от синонимното множество eng-30-03091374-n: {връзка, съединител, конектор}, 'всяко от разнообразни средства в уреди, машини, конструкции, системи и др., служещи за устойчиво скрепяване на елементи, компоненти, устройства, за осъществяване на контакт между отделни части и др.'.

По-специфични зависимости се наблюдават при глаголите *вързвам*, *залепям*, при които **Конектор** са съответно eng-30-03670849-n: {въже}; eng-30-14702416-n: {адхезив}, а при глагола *присаждам* **Конектор** е eng-30-03418242-n: {градински инструмент}.

⁷ Всички примери в текста са извлечени от Българския национален корпус: <http://search.dcl.bas.bg/>.

При някои от глаголите е възможна т. нар. локативна алтернация (Levin 1993: 2), при която **Конекторът** и **Целта** могат да разменят прякообектната и непрякообектната си позиция, както е в примери (1.б.) – (1.д.) и в примери (1.в.) – (1.е.)

1.д. [Денис]_{Агент} *искаше да СВЪРЖЕ* [дървената гривна]_{Конектор} [към двете тръби]_{Обект}.

1.е. [Той]_{Агент} *ВЪРЗА* [връв]_{Конектор} [за разваления зъб]_{Обект}.

3.1.2. Семантичен фрейм Отделяне

Дефиниция на семантичния фрейм: „Агент отделя **Обект** от **Източник** или от друг **Обект**“

Фреймови елементи: Агент, Обект, Източник

Глаголи: коренът на поддървото в Уърднет eng30-01298668-v: {махам} ‘правя така, че да се прекъсне целостта, връзката между две неща или между частите на едно и също нещо’ и хипонимите: {прекъръшвам; прекъръша; скъръшвам; скъръша; откъръшвам; откъръша; счупвам; счупя; отчупвам; отчупя; къртя; откърътвам; откърътя} ‘отделям със сила част от цялото’; {освобождавам; освободя} ‘разкопчавам или разпускам някакъв ограничител’; {откачвам; откача; откачам} ‘отстранявам нещо поставено да виси на закачалка, кука и под. от мястото му’, и коренът на поддървото в Уърднет eng-30-01556921-v: {разделям; разделя} ‘карам две или повече неща или части от нещо да се преместят, отдалечат едно от друго’ и хипонимите му {прекъсвам; прекъсна; откъсвам; откъсна} ‘разделям нещо с късане, правя на късове, нарушавам целостта на нещо’; {откъсвам; откъсна; скъсвам; скъсам} ‘отделям със сила една част от друга част на нещо’.

Примери:

2.а. [Младата жена]_{Агент} *успя само да ОТЧУПИ* [парченца]_{Обект} [от сталактита до главата му]_{Източник}.

2.б. []_{Агент} *ОТКАЧИ* [от колана]_{Източник} [празната ножница]_{Обект} и я метна настрани.

2.в. [Бабичката]_{Агент} *СКЪСАЛА* [от ризата си]_{Източник} [конец]_{Обект}.

2.г. [Пенкроф]_{Агент} *ОТКЪРШИ* [от първото дърво]_{Източник} [два здрави клона]_{Обект}.

Валентен модел:

| Агент | Обект | Източник/Обект |
|-------------|-----------------|----------------------|
| NP (субект) | NP (пряк обект) | PP от (непряк обект) |

Тук прякообектната позиция на **Обекта** и непрякообектната позиция на **Източника** са задължителни.

Селективни особености:

При този тип глаголи могат да се посочат само най-общи селективни специфики. Фреймовият елемент **Агент** се реализира посредством съществителни, хипоними на синонимното множество от Уърднет eng-30-00007846-n: {човек}; докато фреймовите елементи **Обект** и **Източник/Обект** съответстват на хипонимите на eng-30-00002684-n: {физически обект} или eng-30-00001740-n: {обект}. Интересна в тази група е взаимната семантична зависимост – метонимия между **Обект** и **Източник**, както е в примери (2.в.) *риза – конец* (1.д.) и в (2.г.) *дърво – клон*.

3.1.3. Семантичен фрейм **Запълване1**

Глаголите от този подфрейм представят контакта, осъществен при покриване на дадена повърхност от някаква субстанция, вещество.

Дефиниция на семантичния фрейм: „Агент покрива **Цел** (повърхност) с **Тема** (обекти, вещество или субстанция)“.

Фреймови елементи: **Агент**, **Тема**, **Цел**

Глаголи: коренът на поддървото в Уърднет eng-30-01332730-v: {покривам; покрива; закривам; закрива} ‘поставям нещо върху нещо друго или върху някого, така че да го скрива или предпазва’ и хипонимите: {заскреждавам; заскрежда} ‘покривам със скреж’; {мажа; измазвам; измажа} ‘покривам повърхност със слой гъсто вещество, което нанасям още влажно’; {тапицирам; облепвам; облепя} ‘покривам вътрешните стени на жилище, обществена сграда и т. н. с тапети, постери, пана и под.’; {мажа; намазвам; намажа; боядисвам; боядисам} ‘покривам с тънък слой боя, метал и т. н.’; {подплатявам; подплатя} ‘покривам вътрешната страна на дреха с плат за по-голяма топлина и здравина’; {облицовам} ‘покривам лицевата страна, фасадата на сграда’; {завивам; завия; увивам; увия, обвивам; обвия; загъвам; загъна} ‘покривам или нагласям нещо като покривало за предпазване’; {запушвам; запуша} ‘покривам или притискам част от тялото си (уши, нос) за предпазване от шум, миризма и под.’; {подвързвам; подвързва} ‘покривам книга с обвивка от някакъв материал с цел предпазване или заздравяване’; {обличам; облека} ‘поставям емайл върху повърхността на зъб, за да го заздравя; слагам коронка’; {маскирам} ‘слагам маска на или покривам с маска’; {намазвам; намажа}, ‘покривам нещо със слой от някакво вещество’; {нанасям; нанеса} ‘покривам нещо с допълнителен пласт от някакъв материал, вещество’; {бетонирам} ‘поставям бетон върху строителна повърхност, за да я заздравя’; {посипвам; посипя}, ‘покривам с нещо ситно или прахообразно’; {покривам; покрива} ‘поставям покритие, настилка (покрив) на връхната част на постройка, сграда’; {измърсявам; изцапвам; оплесквам; изцапам; измърся; оплескам} ‘покривам някаква повърхност с вещество, което я мърси’; {омазнявам; омазня; намазнявам; намазня} ‘покривам повърхност или материал с някакъв вид масло, обикновено чрез натриване’; {навосъчавам; навосъча} ‘покривам, напоявам, натърквам или излъсквам някаква повърхност с восък’; {напръсквам; напръскам} ‘покривам някакво място изцяло

или частично с пръски течност'; {облепвам; облепя; облицовам}, 'покривам нещо – стена, под, покрив, сграда, с някакво покритие – плочки, керемиди, дървени плоскости и под.'; {забелвам; забеля} 'прикривам, заличавам част от текст с коректор'; {гипсирам} 'покривам и закрепям с гипс, обикновено във вид на превръзка върху счупен крайник'; {затревявам; затревя} 'засявам земна площ с трева или оставям пространство или повърхност да се покрие с трева'.

Примери:

- 3.а. [Тя]_{Агент} *ПОКРИ* почти изцяло [лицето си]_{Цел} [с качулката на наметалото]_{Тема}.
 3.б. [_{Агент} *Замислено НАМАЗА* [кифлата]_{Цел} [с мармалад]_{Тема}.
 3.в. [Христофор]_{Агент} *НАМАЗА* [хайвер]_{Тема} [върху филия хляб]_{Цел}.
 3.г. [Д-р Любенов]_{Агент} *ГИПСИРА* [крака ми]_{Цел} вчера следобед.
 3.д. [_{Агент} *ПОСИПАХА* [подиума]_{Цел} [с дървени стърготини]_{Тема}.
 3.е. [_{Агент} *При езерото ШЯХ ДА НАПРЪСКАМ* [коня си]_{Цел}, за да го охладя.

Валентен модел:

| Агент | Тема | Цел |
|-------------|--|---|
| NP (субект) | NP (пряк обект) PP с/със (непряк обект) | NP (пряк обект) PP върху/на (непряк обект) |

Селективни особености:

При някои глаголи фреймовият елемент **Тема** изразява субстанцията, която се полага на повърхността, и при конкретните глаголи има допълнителни селективни предпочитания при изразяването, например при глаголи като *лакирам* **Агент** е човешко същество, обикновено с някаква професионална квалификация, докато **Тема** е определен вид субстанция, представена в синонимното множество eng-30-04521987-n: {лак}, а **Цел** е eng-30-00002684-n: {физически обект}, eng-30-08660339-n: {повърхност}. Подобни са глаголите *гипсирам*, *затревявам*, *забелвам*, при които фреймовият елемент **Тема** е имплицирана в значението на глагола, обикновено чрез общия корен, и често се изпуска в изречението.

Тук също при някои от глаголите е възможна локативна алтернация, както е в примери (3.в.) – (3.ж.) и в примери (3.д.) – (3.з.).

- 3.ж. [Христофор]_{Агент} *НАМАЗА* [филия хляб]_{Цел} [с хайвер]_{Тема},
 3.з. [_{Агент} *ПОСИПАХА* [дървени стърготини]_{Тема} [върху подиума]_{Цел},

3.1.4. Семантичен фрейм: Изпразване1

Тези глаголи представят контакта, осъществен при отстраняването на някаква субстанция, вещество от дадена повърхност.

Дефиниция на фрейма: „Агент отстранява Тема от повърхността на Източник“

Фреймови елементи: Агент, Тема, Източник

Глаголи: коренът на поддървото в Уърднет eng-30-01532589-v: изчиствам ‘права така, че някой или нещо да стане чист(о), като отстранявам мръсотия или нежелани частици от него’ и неговите хипоними: eng-30-01243982-v: {четкам; изчетквам; изчеткам} ‘изчиствам нещо с четка или по подобен начин’; eng-30-01393996-v: {мета; измитам; измета; помитам; помета} ‘изчиствам помещение, повърхност или друго пространство с помощта на метла, четка или по подобен начин’; eng-30-01535246-v: {пера; изпирам; изпера; мия; измивам; измия} ‘премахвам замърсяване, като го обливам или потапям във вода и почистващ препарат’.

Примери:

4.а. [*Жената*]_{Агент} *ИЗЧИСТИ* [*трохите*]_{Тема} [*от масата*]_{Цел}.

4.б. [_{Агент}] *ИЗМЕТОХ* [*сламата*]_{Тема} [*от пода на кухнята*]_{Цел}.

4.в. *И* [_{Агент}] *да ИЗЧЕТКАМ* [*всичките си сака*]_{Цел} [*с четка за дрехи*]_{Инструмент}.

Валентен модел:

| Агент | Тема | Цел |
|-------------|--|--|
| NP (субект) | NP (пряк обект) PP <i>от</i> (непряк обект) | NP (пряк обект) PP <i>от</i> (непряк обект) |

Синтактичната реализация на някои от глаголите и техните фреймови елементи позволява варианти, при които фреймовите елементи **Тема** и **Цел** могат да разменят прякообектната си позиция, както е при глагола *изчиства* в примери (4.а.) – (4.г.) и *измета* в примери (4.б.) – (4.д.).

4.г. [*Жената*]_{Агент} *ИЗЧИСТИ* [*от трохите*]_{Тема} [*масата*]_{Цел}

4.д. [_{Агент}] *ИЗМЕТОХ* [*пода на кухнята*]_{Цел} [*от сламата*]_{Тема}

Селективни особености:

При този тип глаголи отново могат да се посочат само най-общи селективни специфики. Фреймовият елемент **Агент** се реализира посредством съществителни, хипоними на синонимното множество от Уърднет eng-30-00007846-n: {човек}; докато фреймовите елементи **Обект** и **Източник/Обект** съответстват на хипонимите на eng-30-00002684-n: {физически обект} или eng-30-00001740-n: {обект}. В много от употребите периферният фреймов елемент **Инструмент** е посредникът на контакта между фреймовия елемент **Тема** и повърхността на фреймовия елемент **Цел**, както е в (4.в.). При някои глаголи езиковата инстанция на фреймовия елемент **Инструмент** се съдържа в корена на глагола, както е в примери (4.б.) и (4.в.).

3.2. Глаголи за контакт с контейнер

Тази група от глаголи представлява взаимодействието на обект или субстанция със съд (контейнер), който може да бъде *кутия, чанта, буркан, съд* или всеки друг обект, използван за задържане или съхранение на предмети. Най-характерният фреймов елемент на тези глаголи е **Контейнер** и уточнява конкретния вместиител, съд, с който се осъществява контактът или в рамките на който се извършва действието. Естеството и характеристиките на съда са семантично обвързани със селективните предпочитания на конкретните глаголи и водят до стесняване на общите семантични типове, приписани на фреймовите елементи. Семантичните фреймове и фреймовите елементи, характерни за това множество от глаголи, както и примери за употреба са описани по-долу.

3.2.1. Семантичен фрейм Запълване²

Дефиниция на семантичния фрейм: „Агент поставя Тема (субстанция, вещество, обект) в Контейнер“

Фреймови елементи: Агент, Тема, Контейнер

Глаголи: eng-30-01488956-v: {пълня; напълвам; напълня; товаря; натоварвам; натоваря} ‘слагам, поставям товар в съд, помещение и т. н. до достигане на капацитета му’; eng-30-01489989-v: {натоварвам; натоваря} ‘запълвам определено място или пространство в превозно средство с вещи, стоки, материали и т. н. за транспортиране’; eng-30-01487718-v: {претоварвам; претоваря; пренатоварвам; пренатоваря} ‘натоварвам превозно средство, товарно животно, контейнер и под. с прекалено обемен или тежък товар’; eng-30-01482449-v: {опаковам; пакетирам} ‘поставям или подреждам вещь(и) или стока в специална обвивка от хартия, картон, пластмаса и др.’.

Повечето от глаголите, участващи в този семантичен фрейм, не са организирани в обща дървовидна структура, а са пръснати като листа на други глаголи.

Примери:

5.а. [Те]_{Агент} *НАПЪЛНИХА* [камиона]_{Контейнер} [с тухлите]_{Тема}.

5.б. [_{Агент} Първо *НАПЪЛНИХА* [багажа]_{Тема} [във фургона на Ник]

_{Контейнер}.

5.в. [*Мъжете*]_{Агент} *НАТОВАРИХА* *припряно* [буретата с катран]_{Тема} [в кошовете]_{Контейнер}.

5.г. [_{Агент} *Побързаха да НАТОВАРЯТ* [торбите]_{Тема} [в една малка кола]_{Контейнер}.

Валентен модел:

| Агент | Контейнер | Тема |
|-------------|--|---|
| NP (субект) | NP (пряк обект) PP <i>в/на</i> (непряк обект) | PP <i>с</i> (непряк обект) NP (пряк обект) |

И тук от примери (5.а), (5.б.) и (5.в.), (5.г.) се вижда, че синтактичната реализация на фреймовите елементи **Тема** и **Контейнер** – пряк/непряк обект – не е определена, но е взаимнозависима. Когато фреймовият елемент **Тема** е пряко допълнение, фреймовият елемент **Контейнер** е непряко допълнение и обратно. Възможна е и локативна алтернация, както в примери (5.а.) – (5.д.); (5.б.) – (5.е.); (5.в.) – (5.ж.); (5.г.) – (5.з.).

5.д. [Те]_{Агент} **НАПЪЛНИХА** [в камиона]_{Контейнер} [тухлите]_{Тема}.

5.е. []_{Агент} **Първо НАПЪЛНИХА** [фургона на Ник]_{Контейнер} [с багажа]

_{Тема}
5.ж. [Мъжете]_{Агент} **НАТОВАРИХА** припряно [кошовете]_{Контейнер} [с буретата с катран]_{Тема}.

5.з. []_{Агент} **Побързаха да НАТОВАРЯТ** [една малка кола]_{Контейнер} [с торбите]_{Тема}.

Селективни особености:

При този тип глаголи отново могат да се посочат само най-общи селективни специфики. Фреймовият елемент **Агент** се реализира посредством съществителни, хипоними на синонимното множество от Уърднет eng-30-00007846-n: {човек}; докато фреймовите елементи **Обект** и **Източник/Обект** съответстват на хипонимите на eng-30-00002684-n: {физически обект} или eng-30-00001740-n: {обект}.

3.2.2. Семантичният фрейм **Изпразване2**

Дефиниция на семантичния фрейм: „Агент отстранява **Тема** от **Контейнер**“

Фреймови елементи: **Агент, Тема, Контейнер**

Глаголи: {изпразня; изпразвам} ‘премахвам всичко или цялото количество от нещо, намиращо се някъде’; {разливам; разлея; изливам; излея; изсипвам; изсипя; разсипвам; разсипя} ‘правя така, че течност да премине от един съд в друг или от съд извън него’; {разсипвам; разсипя; разпръсквам; разпръсна} ‘за твърдо вещество, което се състои от дребни частици – карам да се разпилее безразборно от съда, в който е съхранявано’; {изсипвам; изсипя; обръщам; обърна; изливам; излея} ‘оставям съдържимото на някакъв съд да падне, да се разсипе или разлее’. Глаголите, участващи в този семантичен фрейм, не са организирани в обща дървовидна структура, а са пръснати като хипоними на други глаголи за контакт, предизвикващи различен семантичен фрейм като {предавам; предам; пренасям; пренеса}, глаголи за движение {местя; премествам; преместя; измествам; изместя} или глаголи за промяна.

Примери:

6.а. [Той]_{Агент} **РАЗСИПА** [ларвите]_{Тема} [от шепата на Флар]_{Контейнер} [върху каменната настилка на обора]_{Дестинация}.

6.б. []_{Агент} **ПРЪСНАХА** [всички бисквити]_{Тема} [от пакета]_{Контейнер} [в калта]_{Дестинация}.

б.в. [_{Агент} *РАЗЛЯХ* [*остатъка*]_{Тема} [*от бутилката*]_{Контейнер} [*в чашата*]_{Дестинация}·
 б.г. [_{Агент} *Съвсем случайно ИЗЛЯ*[*чая*]_{Тема} [*от канчето*]_{Контейнер}·

Валентен модел:

| Агент | Контейнер | Тема |
|-------------|---|-----------------|
| NP (субект) | PP <i>от</i> (непряк обект/ несъгласувано определение) | NP (пряк обект) |

Примерите показват, че синтактичната реализация на фреймовият елемент **Тема** е пряко допълнение, докато фреймовият елемент **Контейнер** може да се реализира като непряко допълнение и като несъгласувано определение към **Тема**, както в примери (б.а.) – (б.д.), и не подлежи на локативна алтернация. В много от употребите се реализира и периферният фреймов елемент **Дестинация**, който е свързан с преместването на съдържимото (**Тема**) от един вместиел (**Контейнер**) в друг вместиел или върху повърхност (**Дестинация**).

б.д. [*Той*]_{Агент} *РАЗСИПА* [*от шепата на Флар*]_{Контейнер} [*ларвите*]_{Тема} [*върху каменната настилка на обора*]_{Дестинация}·

Селективни особености:

И при този тип глаголи могат да се посочат само най-общи селективни специфики. Фреймовият елемент **Агент** се реализира посредством съществителни, хипоними на синонимното множество от Уърднет eng-30-00007846-n: {човек}; докато фреймовите елементи **Тема** и **Контейнер** съответстват на хипонимите на eng-30-00002684-n: {физически обект} или eng-30-00001740-n: {обект}.

3.3. Глаголи за телесен контакт

В рамките на тази подгрупа глаголите представят действия, изразяващи различни начини, по които повърхността на обект и повърхността на част на тялото влизат в контакт една с друга. Този подклас включва глаголите, представени във Фреймнет със семантичните фреймове **Манипулация** и **Засягане**, които описват „въздействието на **Агент** върху някакъв **Обект**“. Най-характерният фреймов елемент на тези глаголи е периферен – повърхност на **Част на тялото на Агент** (**Body part of agent**), където се осъществява контактът или по която се извършва движението. При семантичния фрейм **Манипулация** фреймовият елемент **Обект** не е дълбоко или постоянно физически засегнат, нито се премества от едно място на друго. Участниците в семантичния фрейм **Засягане** са **Засегнат** и **Засягащ**, като **Засегнат** обикновено е обект на насилствено въздействие.

Въз основа на начина на контакт могат да означават **пряк физически контакт** – {докосвам; докосна; докосвам се; докосна се; допирам; допра; допирам се; допра се} ‘осъществявам физически контакт с някого или нещо’;

{държа} ‘улавям или поемам нещо или някого с ръка или в ръка (ръце) или в хватка’; {стискам; стисвам; стисна} ‘хващам, захващам здраво, с всички сила’; **леко и нежно докосване** – {галя; погалвам; погала; милвам; помилвам} ‘докосвам леко (част от тялото на) някого или нещо в израз на нежни чувства’; {гъделичкам; погъделичквам; погъделичкам} ‘докосвам леко някого или нещо, като предизвиквам лек, приятен сърбеж’; **силно или агресивно докосване** – {удрям; ударя} ‘осъществявам внезапен, често силов контакт между едно тяло (напр. ръка, юмрук, крак, оръжие и т. н.) и друго тяло или повърхност’; {удрям; ударя; бия} ‘рязко и със сила влизам в контакт и упражнявам физическо въздействие върху някого или нещо с част от тялото или с помощта на инструмент’; {щипя; ощипвам; ощипя}, ‘стискам силно между пръстите си’; грижовно или нежно докосване: {прегръщам; прегърна} ‘притискам някого с ръце към себе си в израз на обич или привързаност’, {целувам} ‘докосвам или притискам с устни устните или друга част от тялото на любим човек’.

Примери:

7.а. [Майката]_{Агент} ДЪРЖИ [тенджерата]_{Обект} [с ръка]_{Част от тялото}.

7.б. [Синът ѝ]_{Агент} [я]_{Обект} ПРЕГЪРНА мълчаливо.

7.в. [Брат ѝ]_{Засягащ} [ме]_{Засегнат} УДАРИ [с юмрук]_{Част от тялото}.

Валентен модел:

| Агент/Засягащ | Обект/Засегнат | Част на тялото на Агент/Засягащ |
|---------------|-----------------|---------------------------------|
| NP (субект) | NP (пряк обект) | PP с/със (непряк обект) |

Селективни особености:

Фреймовият елемент **Агент** се реализира посредством съществителни, хипоними на синонимното множество от Уърднет eng-30-00007846-n: {човек}; докато фреймовият елемент **Обект** съответства на хипонимите на eng-30-00002684-n: {физически обект} или eng-30-00001740-n: {обект}, но и на eng-30-00007846-n: {човек}.

При на глаголни синонимни множества, на които е приписан семантичният фрейм **Манипулация**, като eng-30-01216670-v: {хващам; хвана; хващам се; хвана се; държа; държа се} ‘улавям или поемам нещо или някого с ръка или в ръка (ръце) или в хватка’ и неговите хипоними: {стискам; стисвам; стисна; удържа; удържам; притискам; притисна; притисвам} ‘хващам нещо или някого и го държа здраво, неподвижно’; {притискам; притисна; притискам се; притисна се; вкопчвам; вкопча; вкопчвам се; вкопча се; прилепя; прилепвам; прилепя се; прилепвам се; притисвам; притисвам се} ‘хващам здраво нещо или някого, обикновено с две ръце, или допирам плътно цялото си тяло към него’; {люлея} ‘държа някого или нещо, поклащайки го нежно и внимателно’; {сключвам; склуча} ‘съединявам ръцете или пръстите си едни с други или около някого или нещо, като го обхващам, или свързвам край-

щата или насрещните части на нещо, така че да обхванат някого или нещо по обиколката му', както и при хипонимите на eng-30-01212572-v: {улавям; улова; хващам; хвана; грабвам; грабна; захващам; захвана} 'вземам нещо и го задържам здраво' се наблюдават по-специфични ограничения върху ядрените фреймови елементи

За някои глаголи фреймовият елемент **Агент** съответства на хипонимите на eng-30-00007846-n: {човек}, например *притискам се* и *улавям*, докато при други глаголи фреймовият елемент **Агент** може да бъде животно (*стиска*, *притискам се*), т.е. да се изрази с хипонимите на eng-30-08660339-n: {животно}, или **Част на тялото**, съответстващ на хипонимите на eng-30-03183080-n: {част от тялото}. Фреймовият елемент **Част от тялото** е свързан със семантиката на глагола, но рядко се експлицира като в пример (7.а.), а по-често остава неизразен както в пример (7.б.). Ограниченията за фреймовия елемент **Обект** също не могат да се проследят последователно в хиперонимното дърво на Уърднет – **Обект** може да бъде одушевен (*люлея*) или неодушевен (*залавям*).

5. Заключение

В тази студия беше предложено описание на семантичните характеристики и селективните ограничения на глаголи за контакт от Уърднет, които изразяват ситуации, представени от семантичните фреймове **Прикрепване**, **Отделяне**, **Запълване**, **Изпразване**, **Манипулация** и **Засягане**. Разглеждат се конкретни синонимни множества от Булнет и се посочват по-специфични селективни предпочитания за конкретни глаголи или групи глаголи.

По отношение на синтактичната реализация прави впечатление, че част от глаголите и техните фреймови елементи, чието значение представя семантичните фреймове **Прикрепване**, **Запълване1** и **Запълване2**, **Изпразване1** и **Изпразване2**, позволяват варианти, при които фреймовите елементи – съответно **Обект** и **Цел**, **Тема** и **Цел** и **Тема** и **Контейнер** – могат да разменят прякообектната и непрякообектната си позиция. Докато при противоположните по значение семантични фреймове **Отделяне**, **Изпразване1** и **Изпразване2** синтактичната реализация на фреймовите елементи **Тема** – пряко допълнение, и на **Контейнер** – непряко допълнение, е определена и не подлежи на алтернация.

Интересна е наблюдаваната взаимна семантична зависимост на фреймовите елементи **Тема** и **Цел** при семантичните фреймове **Запълване1** и **Запълване2** и фреймовите елементи **Обект** и **Цел** при **Прикрепване**, която заслужава допълнително проучване и може да послужи за допълнително специфициране на селективните предпочитания.

В работата се правят изводи за вътрешната семантична организация на глаголите със семантичен признак *физически контакт*. Представена е част от процеса на подбор, систематизиране и класифициране на глаголите за

контакт в подкласове, отразяващи семантичната и синтактичната съчетаемост на глаголите.

Благодарности: Изследването е извършено по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания – 2020 г.“, Договор No. КП-06-Н50/1 от 2020 г.

Цитирана литература / References

- Джонова 2008: Джонова, М. Аргументна структура на изреченията, съдържащи предикат с непряко допълнение експеримент от типа *спи ми се*. – В: *Bulgarica – studia et argumenta. Юбилеен сборник в чест на 65-та година на проф. дфн Руселина Ницолова*. Мюнхен: Biblion Verlag. (Dzhonova, M. Argumentna struktura na izrecheniyata, sadarzhazhti predikat s nepryako dopalnenie eksperientser ot tipa spi mi se. – In: *Bulgarica – studia et argumenta. Yubileen sbornik v chest na 65-ta godishnina na prof. dfn Ruselina Nitsolova*. Munich: Biblion Verlag.)
- Димитрова-Вълчанова, Декова 2009: Димитрова-Вълчанова, М., Р. Декова. За кодирането на лексикална информация: събития и тяхната лексикализация в английски и български език. – *Български език LVI*, с. 84 – 96. (Dimitrova-Valchanova, M., R. Dekova. Za kodiraniето na leksikalna informatsiya: sabitiya i tyahnata leksikalizatsiya v angliyski i balgarski ezik. – *Balgarski ezik LVI*, pp. 84 – 96.)
- Коева 2021: Коева, С. Предикати за състояние: свойства и лингвистични тестове. – *Zeszyty Cyrylo-Metodiackskie*, с. 106 – 138. (Koeva 2021b: Koeva, S. Predikati za sastoyanie: svoystva i lingvistichni testove. – *Zeszyty Cyrylo-Metodiackskie*, pp. 106 – 138.)
- Костова 2010: Костова, Н. *Основни глаголи за движение в българския език*. София: Авангард Прима. (Kostova, N. *Osnovni glagoli za dvizhenie v balgarskiya ezik*. Sofia: Avangard Prima.)
- Лесева, Стоянова 2021: Лесева, Св., Стоянова, Ив. Семантично описание на глаголи за промяна и йерархична организация на концептуалните фреймове. – В: Коева, С., М. Стаменов (ред.). Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“, София: АИ Проф. Марин Дринов, с. 76-85. (Leseva, Sv., Stoyanova, Iv. Semantichno opisanie na glagoli za promyana i ierarhichna organizatsiya na kontseptualnite freimove – In: Koeva., S., M. Stamenov (Eds.). *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik „Prof. Lyubomir Andreychin“*. Sofia: AI Prof. Marin Drinov, pp. 76-85.)
- Стаменов 2021: Стаменов, М. Групи глаголи за вътрешни психически преживявания. Опит за класификация. – В: Коева, С., М. Стаменов (ред.). Доклади от Международната годишна конференция на Института за български език „Проф. Любомир Андрейчин“. Т. II. София: АИ Проф. Марин Дринов, с. 69 – 75. (Stamenov, M. Grupi glagoli za vatreshni psihicheski prezhivyavania. Orit za klasifikatsia. – In: Koeva., S., M. Stamenov (Eds.). *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik „Prof. Lyubomir Andreychin“*. T. II. Sofia: AI Prof. Marin Drinov, pp. 69 – 75.)

- Тишева 2021: Тишева, Й. Наблюдения върху синтактична реализация на предикати за емоционални състояния. – *Zeszyty Cyrylo-Metodiackskie*, с. 114 – 136. (Tisheva, Y. Nablyudenia varhu sintaktichnata realizatsia na predikati za emotsionalni sastoyania. – *Zeszyty Cyrylo-Metodiackskie*, pp. 114 – 136).
- Baker et al. 1998: Baker, C. F., C. J. Fillmore, J. B. Lowe. The Berkeley Framenet Project. – In: *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, (COLING-ACL'98)*. Montreal, Canada: Association for Computational Linguistics, pp. 86 – 90.
- Chomsky 1965: Chomsky, N. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Dekova 2006: Dekova, R. *Lexical encoding of verbs in English and Bulgarian*. Collections. Institutt for sprek og litteratur [2404]. Doctoral thesis. Det historisk-filosofiske fakultet. <<http://hdl.handle.net/11250/243431>> [01.02.2024]
- Dekova 2011: Dekova, R. *Lexical Encoding of Verbs in English and Bulgarian. An empirical study of lexically encoded information and its formal representation*. LAP LAMBERT Academic Publishing AG & Co. KG.
- Fellbaum 1990: Fellbaum, C. English verbs as a semantic net. – *International Journal of Lexicography*, vol. 3, no. 4, pp. 278 – 301.
- Fellbaum 1999: Fellbaum, C. The organization of verbs and verb concepts in a semantic net. – *Text, Speech and Language Technology*, vol. 6, pp. 278 – 301.
- Fillmore 1970: Fillmore, C. The grammar of hitting and breaking. – In: R. A. Jacobs, P. A. Rosenbaum (Eds.). *Readings in English Transformational Grammar*. Waltham, MA: Ginn, pp. 120 – 133.
- Fillmore 1972: Fillmore, C. The case for case reopened. – In: P. Cole and J. M. Sadock, (Eds.). *Syntax and Semantics: Grammatical Relations*, vol. 8, pp. 59 – 81.
- Fillmore 1982: Fillmore, C. Frame semantics. – In: *Linguistics in the Morning Calm*. The Linguistic Society of Korea, Seoul: Hanshin, pp. 111 – 137.
- Fillmore 2007: Fillmore, C. Valency issues in FrameNet. In T. Herbst, & K. Gutz-Votteler (Eds.), *Valency: Theoretical, descriptive and cognitive issues*. New York: Mouton de Gruyter, pp. 129-160
- Gao 2001: Gao, H. Notions of motion and contact for physical contact verbs. – In: Holmer A., J. Svantesson, E. Viberg (Eds.). *Proceedings of the 18th Scandinavian Conference of Linguistics*, vol. 2, pp. 193 – 209.
- Gao, Cheng 2003: Gao, H., C. Cheng. Verbs of contact by impact in English and their equivalents in Mandarin Chinese. – *Language and Linguistics*, vol. 4, no. 3, pp. 485 – 508.
- Hellan, Dimitrova-Vulchanova 1995: Hellan, L., M. Dimitrova-Vulchanova. *The sign model*, manuscript monograph, Trondheim University.
- Jackendoff 1997: Jackendoff, Ray. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.
- Juffs 1996: Juffs, Al. *Learnability and the lexicon. theories and second language acquisition research*. Amsterdam: John Benjamin.
- Koeva 2021a: Koeva, S. The Bulgarian WordNet: Structure and specific features. – *Papers of BAS*, vol. 8, no. 1, pp. 47 – 70.
- Koeva 2021b: Koeva, S. Towards Expanding WordNet with Conceptual Frames. – In: Sonja Bosch, Christiane Fellbaum, Marissa Griesel, Alexandre Rademaker and Piek Vossen (eds.) *Proceedings of the 11th Global Wordnet Conference*, South African

- Centre for Digital Language Resources (SADiLaR), Potchefstroom, South Africa, 2021, pp. 181 – 191.
- Koeva 2022: Koeva, Svetla. The system of diatheses in Bulgarian and stative predicates. – In: Svetla Koeva, Elena Ivanova, Yovka Tisheva & Anton Zimmerling (Eds.) *Ontology of Stative Situations – Linguistic Modeling. A Contrastive Bulgarian- Russian Study*, Sofia: Prof. Marin Drinov Publishing House of the Bulgarian Academy of Sciences, pp. 117–160.
- Kukova 2020: Kukova, H. Verbs for communication, frame elements and semantic restrictions (on BulNet synsets). – In: *Proceedings of the International Annual Conference of the Institute for Bulgarian Language (Sofia, 2020). Vol. 2*. Sofia: Prof. Marin Drinov Publishing House of Bulgarian Academy of Sciences, pp. 233 – 241.
- Leseva et al. 2019: Leseva, S., I. Stoyanova, M. Todorova, H. Kukova. A theoretical overview of conceptual frames and semantic restrictions on frame elements. – *Linguistique Balkanique LVIII*, no. 2, pp. 172 – 186.
- Leseva et al. 2020: Leseva S., M. Todorova, I. Stoyanova, H. Kukova. A semantic description of the combinability between verbs and nouns (on material from Bulgarian and English). – *Chuzdoezikovo obuchenie (Foreign Language Teaching)*, vol. 47, no. 2, pp. 115 – 128.
- Levin 1993: Levin, B. *English verb classes and alternations: a preliminary investigation*. Chicago: The University of Chicago Press.
- Levin 2015: Levin, B. Verb classes within and across languages. – In: A. L. Malchukov, B. Comrie (Eds.). *Valency classes in the world's languages 2*. Berlin, Germany; Boston, Massachusetts: De Gruyter Mouton, pp. 1627 – 1670.
- Љубнер 2011: Љубнер, S. Concept types and determination. – *Journal of Semantics*, vol. 28, no. 3, pp. 279 – 333.
- Miller 1995: Miller, G. Wordnet: a lexical database for English. – *Communications of the ACM*, vol. 38, no. 11, pp. 39 – 41.
- Miller et al. 1990: Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross, Katherine J. Miller, Introduction to WordNet: An On-line Lexical Database. – *International Journal of Lexicography*, vol. 3, no. 4, pp. 235 – 244
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: extended theory and practice*. Berkeley, California: International Computer Science Institute.
- Vendler 1967: Vendler, Z. *Linguistics in Philosophy*. Ithaca, N. Y.: Cornell Univ. Press.
- Viberg 2004: Viberg, E. Physical contact verbs in English and Swedish from the perspective of crosslinguistic lexicology. – *Language and Computers*, pp. 327 – 352.

ON THE SEMANTIC CLASS OF VERBS OF CONTACT AND DISPLACEMENT

Maria A. Todorova

Institute for Bulgarian Language

maria@dcl.bas.bg

Abstract. The study offers a description of a group of verbs from WordNet with the semantic class (primitive) “verbs for contact”, which belong to the set of commonly used vocabulary in Bulgarian. The description focuses on verbs of contact combined with displacement, which include verbs for attaching, detaching, placing, removing, filling, and emptying. A semantic-syntactic description is proposed, which covers the frame elements of the corresponding FrameNet semantic frame and their corresponding syntactic implementations. Situations represented by the semantic frames **Attachment**, **Detachment**, **Filling**, **Emptying**, **Manipulation** and **Touching** are considered. The analyzed information is based on manual verification, editing and addition of the semi-automatic correlation of two large semantic language resources – WordNet and FrameNet. Subclasses of motion contact verbs have been proposed according to the location of physical contact.

Keywords: *verbs of contact and displacement; semantic-syntactic description*

Maria A. Todorova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52 Shipchenski prohod Blvd., Bl. 17
Sofia 1113 Bulgaria
maria@dcl.bas.bg
<https://orcid.org/0000-0001-5866-7180>

ОСНОВНИ ФРЕЙМОВЕ ЗА ПЕРЦЕПЦИЯ И КОМПЛЕКСНО ПРЕДСТАВЯНЕ НА СЕМАНТИКО-СИНТАКТИЧНА ИНФОРМАЦИЯ

Христина Кукова

Институт за български език „Проф. Любомир Андрейчин“,
Българска академия на науките

hristina@dcl.bas.bg

Резюме. Студията е посветена на перцептивни глаголи и концептуалното им описание в българския език. Направен е аналитичен преглед на досегашни теоретични обобщения, както и на научноприложни изследвания. Като отправна точка за настоящото проучване служи разнообразието от конструкции и синтактични реализации, които представителите на този клас предикати проектират върху изреченското си обкръжение. Анализът се основава на семантична и отчасти морфологична информация за лексикалните единици от Уърднет (Fellbaum 1998; Коева 2021a), както и на семантико-синтактични специфики, с които изследваните глаголни единици са представени във Фреймнет (Baker et al. 1998; Ruppenhofer et al. 2016).

Разгледани са 7 семантични фрейма¹, които проявяват висока фреквентност и и се извикват от голямо разнообразие от глаголни единици. Описанието включва граматични особености на глаголите, семантични и синтактични ограничения, които глаголните единици налагат върху фреймовите елементи, съотнасяне на фреймовите елементи с определено синонимно множество или семантичен клас от съществителни имена в Уърднет, отразяващи реализацията на елементите от фрейма. От Фреймнет е заимстван статусът на фреймовите елементи, който е от съществена важност за реализацията на дадена лексикална единица.

Ключови думи: глаголи за перцепция, семантични фреймове, фреймови елементи

* Христина Кукова. Основни фреймове за перцепция и комплексно представяне на семантико-синтактична информация. – В: Св. Коева (ред.) *Семантични студии. Синтактично и семантично описание на български глаголи*. София: Издателство на БАН „Проф. Марин Дринов“, с. 154 – 178.
<https://doi.org/10.7546/SemanticStudies2024.07>

¹ Описани са наблюденията върху 6 семантични фрейма и прототипния фрейм **Перцепция (Perception)**, който обобщава останалите.

1. Въведение

Студията си поставя за цел да онагледя употребата на глаголи за перцепция, които принадлежат към основния речников състав в българския език. Приложена е методологията на фреймовата семантика, за да се открият разнообразните конструкции, които изграждат глаголите за перцепция. За описанието на отделните значения е използвана лексикално-семантичната мрежа Булнет. Така представените значения на изследваните глаголи са допълнени с лексикална, морфологична и фреймова информация за предикатите, както и данни за селективните ограничения на техните фреймови елементи под формата на специфични синонимни множества от Уърднет. В повечето случаи подредните понятия (хипоними) на избраните синонимни множества могат да заемат позицията на съответния фреймов елемент.

Изследването е корпусно ориентирано и представя данни, ясно отразяващи важноста и спецификата на контекста при семантико-синтактичния анализ. Същинската работа върху глаголите се основава на авторитетното изследване на Б. Левин (Levin 1993), което демонстрира взаимовръзката между компонентите на значението на глагола, синтактичните му свойства и естеството на аргументите му. Левин разпределя глаголите за възприятие в 4 големи подкласа на основата на семантично групиране, както и на участието на предикатите във валентни алтернативи. Авторката класифицира глаголите в групите 1) виждам (see), 2) забелязвам (sight), 3) надничам (peer) и 4) стимул-субектни глаголи (Levin 1993: 185 – 188).

Синтактичните различия обичайно могат да бъдат предсказани от лексикалната семантика, затова е отделено специално внимание на взаимодействието между лексикално-семантичните параметри и синтактичното поведение както на предиката, така и на фреймовите елементи. Проучването на корпусните данни дава възможност последователно и непротиворечиво да се опишат теоретичните и емпиричните наблюдения върху езиковия материал.

Останалата част от студията е организирана, както следва: Част 2. е посветена на представяне на използваните ресурси, както и на понятията за семантичен и концептуален фрейм. В Част 3. са очертани досегашни изследвания и мотивацията, която стои зад проучването. В нея са представени методологията на работа, описание на класа на перцептивните предикати и различни критерии за класификацията им. Разнородните подходи при описанието и таксономията на глаголите се използват като основа за диференциация на подкласовете в рамките на цялостното изследване. В Част 4. е концентриран основният аналитичен преглед на изследваните глаголи и техните специфики в българския език. Основната част е съсредоточена върху семантико-синтактичните свойства на глаголите за перцепция и техните семантични фреймове, както и на синтактичното поведение на фреймовите елементи. Направен е опит да се систематизират възможните семантични ограничения в рамките на подкласовете. Част 5. представя обобщение на резултатите и насоки за бъдеща работа.

2. Ресурси

Уърднет е основният ресурс, използван за анализ на лексикално-семантичните и отчасти на морфологичните свойства на предикатите. Уърднет е лексикално-семантична мрежа, подходяща за машинна обработка, която води началото си от Принстънския университет и е разработена от екип, ръководен от Дж. Милър (Miller 1995; Fellbaum 1998). Българската му версия Булнет съдържа над 110 000 единици, наречени синонимни множества или синсети (Коева 2021а).

За изследване и представяне на семантико-синтактичната комбинаторика на предикатите в рамките на настоящото проучване е използван Фреймнет (Ruppenhofer et al. 2016). Фреймнет е ресурс, създаден под ръководството на Ч. Филмор през 1997 г. (Baker et al. 1998), и е от съществено значение както за теоретичните лингвистични изследвания, така и за практическата обработка на естествен език. Семантичните фреймове представляват концептуална структура на дадено събитие или обект, а неговите участници са дефинирани като фреймови елементи. Те биват ядрени и неядрени, като първите са от съществено значение за реализирането на всеки уникален фрейм, докато вторите са предимно описателни (за време, място и т. н.) (Ruppenhofer et al. 2016: 23). Всяко значение на една и съща дума се кодира като отделна лексикална единица, която се асоциира с различен семантичен фрейм.

Както посочва С. Коева, концептуалният фрейм съдържа уникален набор от синтагматични отношения между глаголен литерал от Булнет и синонимни множества от съществителни имена, които могат да заемат позиции на фреймовите елементи (Коева 2021b: 182). Въз основа на това, което е известно за глаголното значение от Уърднет, и на синтактичните и семантичните ограничения, които то налага на своите фреймови елементи от Фреймнет, генерираме мрежа от възможни комбинации.

С оглед на изчерпателното описание на българските глаголни лексикални единици и обкръжението им е използвана информация от двата ресурса, както следва:

а) от Фреймнет: посочват се ядрените фреймови елементи с техните семантични типове от съответните семантични фреймове, подходящи множества от глаголни лексикални единици, свързани със специфичен семантичен фрейм, и основните релации, които го характеризират;

б) от Уърднет: лексикалните значения; хиперо-хипонимните отношения, които организират именните и глаголните синонимни множества в дървета и поддървета; семантичните класове, към които принадлежат.

Глаголите в рамките на изследването са разгледани в конкретно значение, а техните фреймови елементи могат да бъдат NP, PP, AdvP, AccCl, DatCl или изреченски комплемент (S или малко изречение). Използваме уеббазираната система БулФрейм за създаване и визуализиране на концептуалните

фреймове (Коева, Douchev 2022: 206 – 210), в която данните са организирани, както следва:

1) Лексикален раздел: лема на глагола, информация дали глаголът е еднословен или съставен, част на речта, идентификационен номер от Уърднет, номер на значението, дефиниция на значението, семантичен клас на синонимното множество, стилистична бележка или бележка за употреба.

2) Морфологичен раздел: морфосинтактичните особености на глагола могат да бъдат от съществено значение при формирането на структурата на изречението. Включени са морфологични характеристики: вид на глагола (несв. в., св. в., *imperfectiva tantum*, *perfectiva tantum*, несв. и св. в.), транзитивност (преходен или непреходен) и др.

3) Фреймов раздел: Една част от компонентите в раздела са наследени от Фреймнет (Ruppenhofer et al. 2016): име на семантичния фрейм, дефиниция на фрейма, междуфреймови релации, фреймови елементи с техните имена и статус (ядрени, неядрени и екстратематични) и дефиниция. Други са специфични за езика и се използват за изразяване на уникалността на синтактичните и семантичните параметри на глагола. Синтактичните фрази, които могат да бъдат реализация на фреймови елементи в българския език, са: NP (съществително име), PP (предложна фраза), AdvP (наречие), S (изречение), SC (малко изречение), AccCL (задължителна акузативна клитика), DatCL (задължителна дателна клитика) (Коева, Douchev 2022: 208).

Анализът е илюстриран с богат емпиричен материал, отразяващ значението на контекста при семантико-синтактичния анализ на семантичните фреймове. Езиковият материал е ексцерпиран от Българския национален корпус (Коева et al. 2012), създаден в Института за български език „Проф. Любомир Андрейчин“.

3. Предишни изследвания и мотивация

3.1. Методология

С. Коева отбелязва необходимостта от описание на синтагматичните отношения в Уърднет, като въвежда абстрактната структура концептуален фрейм, в която множество от глаголни лексикални единици демонстрира еднозначни синтагматични отношения с подходящи множества от съществителните (Коева 2021b: 182).

С. Лесева и др. изследват възможността за интегриране на данни от Уърднет, Фреймнет и Върбнет, като предлагат „система от семантични отношения, отразяващи тематичните връзки между предикатите и техните потенциални аргументи в рамките на Уърднет“ (Лесева и др. / Leseva et al. 2018).

Нашият подход, основан на фреймовата семантика (Fillmore 2002; Baker et al. 1998; Baker et al. 2003; Ruppenhofer et al. 2016; Коева / Коева 2010), съчетава както синтактичния абстрактен методологичен апарат, така и проекцията на семантичните и прагматичните отношения върху възможните фреймови елементи. В контекста на своята специфика перцептивните глаголи представляват интересно поле за мултидисциплинарни изследвания.

3.2. Класификация на глаголите за перцепция

Класът на перцептивните предикати е обединен от семантичния елемент сетивно възприятие. С. Моисеева ги характеризира като лексикално-семантично поле от множество езикови единици, които се групират около ядрен семантичен компонент (Моисеева / Moiseeva 2005: 9). Тъй като глаголите за възприятие отразяват връзката между физиологичната перцепция и лингвистичното изразяване, основните подкласове, на които се делят тези глаголи, са обвързани с петте вида сетивно възприятие в зависимост от участието на определен сетивен орган – зрително, слухово, осезателно, обонятелно и вкусово (Viberg 1983; Вежбицкая / Vezhbickaja 1986: 338).

Представителите на всеки един от тези подкласове описват ситуации, в които експериментът (възприемащият участник в ситуацията) ментално обработва даден външен стимул (обект или ситуация) с помощта на сетивен орган (Verhoeven 2007: 49). Глаголите за възприятие могат да присъединяват в качеството на изреченски субект както възприемащия, така и стимула. Ще наричаме първия подклас глаголи със субект-експеримент (виждам, чувам), а втория – със субект-стимул (изглеждам, звучи). Субект-експериментните глаголи, от своя страна, могат да описват две разновидности на възприятието – волева (*слушам*) и пасивна (*чувам*) перцепция, вж. (Viberg 2001: 1296). Следвайки пресечните точки на семантичната и синтактичната подялба, можем да онагледим всички подкласове перцептивни предикати в Таблица 1. (Кукова / Kukova 2022: 24).

| Сетива | Субект-експериментни | | Обект-експериментни |
|------------|----------------------|----------------|---------------------|
| | Волеви | Спонтанни | |
| зрително | <i>гледам</i> | <i>виждам</i> | <i>изглеждам</i> |
| слухово | <i>слушам</i> | <i>чувам</i> | <i>звуча</i> |
| тактилно | <i>опитвам</i> | <i>усещам</i> | <i>на допир съм</i> |
| олфактивно | <i>душа</i> | <i>надуша</i> | <i>мириша</i> |
| вкусово | <i>опитвам</i> | <i>вкусвам</i> | <i>имам вкус</i> |

Таблица 1. Типове предикати за перцепция

Глаголите от първата колона (волеви субект-експериенцерни предикати) ще бъдат разгледани в Част 4.2. в рамките на семантичния фрейм **Активна перцепция (Perception active)**, а от втората колона (спонтанни субект-експериенцерни предикати) – на фрейма **Пасивна перцепция (Perception experience)** в Част 4.3. Обект-експериенцерните предикати за перцепция ще бъдат анализирани с помощта на семантичните фреймове в Част 4.4., 4.5. и 4.6. – съответно **Перцепция посредством тялото (Perception body)** (*боли, боде*), **Създаване на шум (Make noise)** (*тракам, кънтя*) и **Шум от движение (Motion noise)** (*бръмча, трополя*). Семантичните фреймове, описващи звукопроизводство, са подбрани като обхващащи най-многобройни и представителни глаголни лексикални единици спрямо останалите подкласове обект-експериенцерни предикати. С цел изчерпателност на описанието и онагледяване на необходимостта от конкретизиране на значението при многозначните глаголи, в Част 4.7. ще бъде отразен и фреймът **Предизвикване на шум (Cause to make noise)** (*звъня, бия*).

3.3. Приложения на изследването

Анализът на семантичните фреймове и фреймовете им елементи може да бъде формализиран и организиран в база от данни, която да допринесе за развитието на много задачи, свързани с компютърната обработка на естествения език, например автоматично приписване на семантични роли, класификация на текстове, машинен превод, извличане на информация и предоставяне на отговори на конкретни въпроси. Такъв анализ играе ключова роля и при корпусните изследвания, подпомагайки разбирането на структурата и семантиката на текста в рамките на даден корпус. Този тип изследвания усъвършенстват семантичната анотация и благоприятстват съпоставителните изследвания в рамките на един или повече езици, жанрове или периоди. Могат да бъдат изведени фините нюанси на контекстуалните значения, което да доведе до непротиворечива идентификация на полисемията, омонимията и метафоричните измествания в рамките на един корпус. Не на последно място, този тип анализ върху големи корпуси може да доведе до компилация на ресурси, състоящи се от семантични фреймове, лексикални единици, които ги пораждат, и фреймови елементи, което води до разширяване и по-задълбочено разбиране на фреймовата семантика.

Езиковоспецифичното разгръщане на фреймовете осигурява богато представяне на значението и структурата на езика, което позволява на моделите за обработка на естествения език да изпълняват по-ефективно широк спектър от задачи. То осигурява също така теоретична и практическа основа за разнообразни корпусни изследвания, като позволява на изследователите да навлязат по-дълбоко в семантиката на текстовете, да сравняват различни корпуси, да откриват нови модели и да подобрят разбирането ни за структурата и значението на езика в рамките на даден корпус.

Фреймовосемантичните изследвания могат да допринесат и за решаване на конкретни теоретични въпроси. Една от нерешените теоретични задачи, която стои пред лингвистите, е дефинирането на ясни селективни ограничения за голям набор от глаголи. Фреймовите елементи във Фреймнет са маркирани с някои общи семантични ограничения – Разумно същество (Sentient), Артефакт (Artifact) и др., но не съществува изчерпателна типология за всички (ядрени) фреймови елементи (Коева 2020). В теоретичния си обзор С. Лесева и др. предлагат дефиниции на селективните ограничения на основа на семантичните характеристики на елементите, като използват информация от Фреймнет и Върбнет (Leseva et al. 2019: 176). Авторите очертават описания на фреймови елементи, които кореспондират със семантични категории в Уърднет и Фреймнет (Човешко същество, Волево, Конкретно, Съдържание, Абстрактно и др.). Тъй като основната задача е да посочим приемливите комбинации от предикати, представени със синонимните си множества, и класове от съществителни, е необходимо да се направи по-подробен анализ на семантиката на участниците в ситуацията и при необходимост те да бъдат обобщени или детайлизирани.

4. Семантични фреймове и фреймови елементи

В тази част от изследването ще разгледаме трите основни и най-многобройни семантични фрейма, свързани с глаголите за т. нар. същинско възприятие – **Активна перцепция, Пасивна перцепция, Перцепция посредством тялото**, обединени от нелексикалния фрейм **Перцепция (Perception)**. Следват три фрейма, обхващащи един конкретен подклас глаголи за звукопроизводство – **Създаване на шум, Шум от движение, Предизвикване на шум (Cause to make noise)**, които добре онагледяват взаимовръзката между семантично сходни, но синтактично диференцирани употреби на глаголите. Основната цел, която е поставена, е да бъде описан за български всеки фрейм с ядрените си фреймови елементи (основните участници в описваната ситуация), като се дефинират селективните им ограничения и съответната синтактична реализация на всеки от фреймовите елементи. В резултат от анализа на семантико-синтактичното обкръжение, характерно за всеки глаголен подклас, са дефинирани по-детайлни семантични ограничения по отношение на възможните именни или предложни групи, с които се комбинират.

Тази комбинация се основава върху концептуалната структура на глаголните синонимни множества. Всяко едно от тях се съпоставя със семантичните фреймове, приписани от Фреймнет, като се проследява дали глаголите отговарят в семантично отношение на конкретната употреба. За фреймовите елементи, които получават синтактична реализация, се прецизират възможните именни синонимни множества от Уърднет.

Като изходни се приемат дефинициите на семантичните фреймовете във Фреймнет. В случаите, в които е необходимо модифициране, фреймовете са представени, като е отбелязана необходимостта от прецизиране за български език.

4.1. Перцепция

Когато става въпрос за глаголи от класа на перцепцията, необходимо е да разгледаме прототипния фрейм **Перцепция**, който обобщава характеристиките за това семантично поле семантични специфики.

Дефиниция във Фреймнет²: „**Възприемащият (Perceiver)** възприема даден **Феномен (Phenomenon)**“. Общият фрейм **Перцепция** се наследява като основен фон при всички фреймове, в които съзнателно одушевено същество реагира на промените в околната среда с помощта на различни сетивни органи. Фреймовете, свързани с релация **Наследяване** или други, могат да специфицират типа сетивност (зрителна, слухова, тактилна, вкусова, обонятелна), да конкретизират актовете на възприятие (*надничам, подслушвам*) или качествата на възприемания **Феномен** (*блестя, тракам*). Ядрени фреймови елементи на този семантичен фрейм са **Възприемащ** и **Феномен**.

Възприемащ – фреймовият елемент **Възприемащ** се отнася до някого, чието когнитивно състояние или усещане (знание, вярване, идея, преценка или класификация) се описва от ситуацията .

Феноменът е същността, която се възприема от **Възприемачия**.

1. [Сър Томас]_{Възприемащ} *забеляза [издутината]*_{Феномен} *и леко побледня.*

4.2. Активна перцепция

Семантичният фрейм се предизвиква от глаголни синонимни множества със семантичен примитив перцептивен глагол, при който е налице волево насочване на вниманието на перципиента. Свързан е с релацията **Наследяване** с нелексикалните фреймове **Умишлено действие** (Intentionally act) и **Перцепция**. Характерни представители са глаголи от групата на активните предикати за възприятие:

eng-30-02130300-v³: {разглеждам; разгледам; поглеждам; погледна; гледам}; eng-30-02125032-v: {душа; подушвам; подуша};

eng-30-02152278-v: {преглеждам; прегледам; прехвърлям; прехвърля; пробягам с поглед; пробягам с поглед; прокарвам поглед; прокарам поглед};

eng-30-02125032-v: {подушвам; подуша; помиривам; помириша; вдъхвам; вдъхна; вдишвам; вдишам}.

² Дефинициите на семантичните фреймове и фреймовите елементи в изследването представляват свободен превод от Фреймнет <<https://framenet.icsi.berkeley.edu/>>.

³ Синонимните множества се индексират с т. нар. междуезиков индекс (InterLingual Index – ILI), с помощта на който са свързани лексикалните съответствия в различните езици.

Предикатите за зрително възприятие многократно превъзхождат в числово отношение останалите подгрупи.

Дефиниция във Фреймнет: „Фреймът съдържа предикати за перцепция, при които перципиентът волево насочва вниманието си към някаква същност с цел сетивно възприятие“. **Възприемащият** е активен, затова ролята му в този фрейм е **Агентивен възприемащ (Perceiver agentive)**.

Ядрените фреймови елементи са **Агентивен възприемащ**, който извършва определена волева дейност, за да получи сетивното възприятие, и **Феномен** – същността, към която възприемащият насочва вниманието си. В изречения в активен залог **Агентивният възприемащ** е изразен като подлог, а **Феноменът** – като допълнение (2.).

2. [Д-р Хелси]_{Агентивен възприемащ} **ПРЕГЛЕДА** [*частично разтопените компютърни компоненти*]_{Феномен}.

Неизразени ядрени елементи в рамките на фрейма могат да бъдат **Част на тялото (Body part)** и **Посока (Direction)**.

Фреймовият елемент **Част на тялото** маркира сетивния орган, използван от **Агентивния възприемащ** при възприятието, за да получи сетивното усещане. Той обикновено е изразен с предложна фраза (3. – 5.). Този фреймов елемент по правило се среща с глаголите за тактилно възприятие, които не посочват експлицитно с коя част от тялото се извършва възприятието (4.).

3. *Стискаше парцалена кукла и ги ГЛЕДАШЕ* [*с големите си очи*]_{Част на тялото}.

4. *Никаква промяна в температурата. Дали да я ПРОБВАМ* [*с пръст*]_{Част на тялото}, *който няма да ми липсва много?*

5. *Източният вятър донесе уханието на топъл хляб от градските фурни, което той долови, а когато минавахме през пристана в Зевгма, ПОДУШИ* [*с носа си*]_{Част на тялото} *и каза: – Усещам мириса на канела, сандалово дърво и моряци.*

По-голямата част от предикатите за възприятие съдържат имплицитно използваната част на тялото и затова този фреймов елемент се появява непоследователно в рамките на семантичния фрейм. Например зрителното възприятие е задължително свързано с органите за зрение, слуховото – с органите за слух и т. н. Макар и в редки случаи, с фреймовия елемент **Част на тялото** може да се специфицира как точно се извършва действието (6. – 7.).

6. *СЛУШАЙ* [*с ушите си*]_{Част на тялото}, *не с устата.*

7. *Уил ГЛЕДАШЕ* [*с широко разтворени очи*]_{Част на тялото}.

Посоката обозначава изрази, описващи как се насочва вниманието на перципиента по време на акта на перцепция, като изключва онези от тях, които назовават **Местоположението на възприемащия (Location of perceiver)**. Обичайно фреймовият елемент **Посока** се проектира като предложна фраза и се въвежда

от предлога *през* (8. – 11.). Среща се в изречения, маркиращи различни подкласове възприятие – зрително (8., 10.), слухово (9.), обонятелно (11.).

8. *Един слуга, който седеше зад него в изправен сандък с форма на ковчег и НАБЛЮДАВАШЕ [през процеп в дъските]_{Посока}, видя сигнала и дръпна силно усуканото въже.*

9. *Досадно е – продължи раздразнена тя, – че човек не може едновременно да гледа и да СЛУША [през ключалката]_{Посока}.*

10. *Изад вратата надзирателят ги СЛЕДЕШЕ [през прозорче от метастъкло]_{Посока}.*

11. *После пак до вратата, пак ДУШЕШЕ [през пролуката]_{Посока} и установи по миризмите: на двора става нещо тревожно.*

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози |
|-----------------------------|--|--|-----------------------|---------------------------------|--------------------|
| Агентивен възприемащ | Лице или друга одушевена същност, която извършва определена волева дейност, за да получи сетивното възприятие. | eng-30-00007846-n: {човек} eng-30-00015388-n: {животно} | NP | Подлог | — |
| Феномен | Същност, към която възприемащият насочва вниманието си. | eng-30-00001930-n: {физическа същност} | NP PP S | Допълнение, Подчинено изречение | В, във, към |
| Част на тялото | Сетивен орган, използван от Агентивния възприемащ при възприятието, за да получи сетивното усещане. | eng-30-05299178-n: {сетивен орган} | NP | Обстоятелствено пояснение | С, със |

| | | | | | |
|--------|--|-----------------------------------|----|---------------------------|-------------------|
| Посока | Място, пространствена точка или последователност от пространствени точки, описващи как се насочва вниманието на перципиента по време на акта на перцепция. | eng-30-02735688-n: {пространство} | NP | Обстоятелствено пояснение | В, във, към, през |
|--------|--|-----------------------------------|----|---------------------------|-------------------|

Таблица 2. Ядрени фреймови елементи на семантичния фрейм **Активна перцепция**

4.3. Пасивна перцепция

Дефиниция във Фреймнет: „Семантичният фрейм описва перцептивни глаголи, чиито възприемащ изпитва акта на възприятие, без задължително да е имал намерение да го направи“. Поради тази причина възприемащият в рамките на този фрейм се нарича **Пасивен възприемащ**. Фреймът е свързан с релацията **Наследяване с Перцепция**, а с релацията **Използван от** – с **Перцепция посредством тялото**.

Типични представители на лексикалните единици от този семантичен фрейм са предикатите за пасивен тип перцепция:

- eng-30-02163746-v: {виждам; вида; забелязвам; забележа};
- eng-30-02129879-v: {зървам; зърна};
- eng-30-02154312-v: {съзирам; съзра; съглеждам; съгледам; регистрирам};
- eng-30-02107442-v: {чувам; чуя};
- eng-30-02189168-v: {дочувам; дочуя; долавям; доловя; подочувам; подочуя};
- eng-30-02106006-v: {усещам; усетя; чувствам; почувствам}
- bg-0000015L: {доживявам; доживея}.

Трите ядрени фреймови елемента са **Пасивен възприемащ (Perceiver passive)**, **Феномен**, **Част на тялото**.

Част на тялото – локализира мястото от тялото на **Пасивния възприемащ**, където се осъществява усещането. Фреймовият елемент е синтактично изразен с предложна фраза (12.).

12. *Том, отдавна свикнал на непрекословно подчинение, взе харпуна си и тръгна тихомълком по стъпките на своя нов началник, но изведнъж УСЕТИ [на рамото си]_{Част на тялото} ръката на лейтенанта.*

И в рамките на този семантичен фрейм неспецифицираните по сетивност типове възприятие присъединяват този фреймов елемент, тъй като възприятието не е локализирано в конкретна част на тялото (13. – 15.). Макар и по-рядко, фреймовият елемент **Част на тялото** се среща и с глаголи, които изразяват конкретен тип възприятие (16. – 17.).

13. *Нещото, което се случи и промени всичко, беше, че УСЕТИХ [с ръката си]_{част на тялото} леко бутване отвътре.*
 14. *Той го ПОЧУВСТВА [с кожата си]_{част на тялото}.*
 15. *Седнете – нареди той на Паул, като помириса и ОПИТА [с небцето си]_{част на тялото} вкуса на възгорчивото бледо вино.*
 16. *Лежах на петнайсетина крачки от него и ЧУХ [с ушите си]_{част на тялото} и ВИДЯХ [с очите си]_{част на тялото} всичко.*
 17. *Един от пленниците го ВИДЯЛ [с очите си]_{част на тялото}.*

Ако с *небцето си* в пример (15.) характеризира конкретна **Част на тялото**, с която се възприема, в сходния пример (18.) предложната фраза маркира неядрения фреймов елемент **Място (Place)**.

18. *Ричър все още УСЕЩАШЕ вкуса на цигарата [в устата си]_{място}.*

В рамките на фреймовия елемент **Част на тялото** могат да бъдат изразени още и метафорични значения, свързани със степента на внимание (19. – 20.) или истинността на възприеманото (21.).

19. *Я ме ЧУЙ [с двете уши]_{част на тялото}, ти, срам за рода си.*
 20. *Умът ѝ набързо подреди това, което бе ЧУЛА [с половин ухо]_{част на тялото} от сърдитото му боботене.*
 21. *И за Зверовете се говореше, че били мит, но те ги бяха ВИДЕЛИ [със собствените си очи]_{част на тялото}.*

При глаголи, които бележат многозначност по оста активно-пасивно възприятие, каквито са повечето глаголи за обоняние, употребата на този фреймов елемент често разграничава двата типа, като се използва предимно със значенията за активно възприятие (22.).

22. *Вуйчо Рубън обаче първо погледна надолу към долината и ПОДУШИ [с нос]_{част на тялото}, като че ли искаше да я помирише.*

Пасивен възприемащ – посочва субекта на възприятието, не задължително волево целящ това. В синтактично отношение обикновено се изразява от субекта (23.).

23. *[Ранд]_{Пасивен възприемащ} УСЕТИ как отърля дрехата му, НАДУШИ обгорената вълна.*

Феномен – този фреймов елемент маркира обекта или стимула, който възприемащият усеща със сетивата си. Обичайно на изреченско ниво заема позицията на пряк обект (24.).

24. *Тръгнахме по тунела, после ВИДЯХМЕ [светлина]_{Феномен} и ЧУХМЕ [гласове]_{Феномен}.*

Фреймът обхваща и лексикални единици, които не са обвързани с конкретен тип възприятие и които можем да категоризираме като общо възприятие (25.).

25. *Искам да изтъкна само, че събитието трябва да има известно условие за съществуване на равнище, което не можем да ДОЛОВИМ с нормалните си възприятия.*

Съгласно йерархията на Виберг⁴ (зрително > слухово > тактилно > вкусово, обонятелно възприятие), по-високо разположените в нея предикати могат да заемат мястото на по-ниско разположените или да се използват при липсата на такива (Кукова / Kukova 2022: 31; Стефанов / Stefanov 2024). В тези случаи наблюдаваме конкретизация посредством типа на **Феномена** или органа на възприятие, а понякога – и на двете. Така можем да заключим, че литералите от синонимното множество eng-30-02110552-v: {усещам; усетя; чувствам; почувствам} се срещат освен за назоваване на тактилно възприятие и по отношение на вкусово и обонятелно възприятие. Глаголните лексикални единици от синонимното множество eng-30-02117955-v: {улавям; уловя; долавям; доловя} покриват всички видове възприятия. В примери (26. – 31.) е онагледена разнородната употреба с един от глаголите, обозначаващ различни видове сетивно възприятие (зрително (26.), слухово (27. – 28.), обонятелно (29. – 30.), вкусово (31.). Значението се конкретизира от контекста с помощта на използвания тип сетивност или стимул.

26. *Кат ДОЛОВИ движение [с периферното си зрение]_{Част на тялото} и рязко се извърна, готов да побегне или да се бие, ако е нужно.*

27. *[Ушите]_{Част на тялото} му ДОЛОВИХА пращането на огън и клокочещ звук, източник неизвестен.*

28. *Смътно ДОЛОВИ [крясъците]_{Феномен}*

29. *Подуши въздуха и ДОЛОВИ [мирис на живот]_{Феномен}*

30. *ДОЛОВИ [с обонянieto си]_{Част на тялото} [миризмата на естери от предпазващо масло]_{Феномен}, оказали се по-силни от мириса на Монеовата пот.*

31. *Дали рецепторите ми не са толкова изтънчени, колкото се предполага, но не ДОЛОВИХ [вкуса на кокоса]_{Феномен} в общото съчетание.*

В корпуса се срещат примери, в които снемането на многозначността се основава както на присъствието на **Частта на тялото** или **Мястото**, така и на **Феномена** (32.).

32. *Млъкна, когато УСЕТИ [на челото си]_{Място} [студената тясна ръка]_{Феномен}*

⁴ Вж. Viberg 1983, 2001.

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози |
|--------------------------------|--|---|--------------------------|---------------------------------------|-------------------------|
| Пасивен възприемащ | Субект на възприятието, не задължително волево целящ това. | eng-30-00007846-n: {човек}; eng-30-00015388-n: {животно} | NP | Подлог | — |
| Феномен | Същност или стимул, който възприемащият възприема със сетивата си. | eng-30-00001930-n: {физическа същност} | NP S | Допълнение, Подчинено изречение | — |
| Част на тялото | Сетивен орган, използван от Пасивния възприемащ при възприятието, за да получи сетивното усещане. | eng-30-05299178-n: {сетивен орган} | PP | Обстоятелствено пояснение | С, със, на, върху |

Таблица 3. Ядрени фреймови елементи на семантичния фрейм
Пасивна перцепция

4.4. Перцепция посредством тялото

Дефиниция във Фреймнет: „Този семантичен фрейм съдържа глаголни лексикални единици, описващи физически усещания, които могат да бъдат породени във всяка част на тялото“. Свързан е с фрейма **Пасивна перцепция** посредством релацията **Използва**.

Характерни представители на тази група глаголни лексикални единици са:

eng-30-02121511-v: {боли ме; боли; имам болка};

eng-30-02198819-v: {убива; убива ми};

eng-30-02122522-v: {обажда се, обади се};

eng-30-02123424-v: {бода};

eng-30-02122665-v: {пулсирам; туптя; тупкам};

eng-30-02121048-v: {паря; горя}.

Ядрените фреймовите елементи са **Част на тялото** и **Експериенцер** (**Experiencer**) и са дефинирани във Фреймнет, както следва:

Част на тялото – назовава мястото върху тялото, където се осъществява физическото усещане. Предикатите традиционно присъединяват съответната **Част на тялото** която се изразява експлицитно в субектна позиция (33.). Възможно е също проектирането на фреймовия елемент като предложна фраза (34.) или пряк обект (35.).

33. *Все още ме БОЛИ [главата]_{Част на тялото}*

34. *Ронин се размърда, опитвайки се да намери по-удобно положение, като се надяваше, че така ще облекчи натиска – и откри, че нещо го БОДЕ [по врата]_{Част на тялото}*

35. *При всяко движение мокрите кичури коса удряха гърба ѝ и сякаш ледени остриета БОДЯХА [кожата ѝ]_{Част на тялото}*

Експериенцер – **Експериенцерът** е одушевената същност, която изпитва физическо усещане – вътрешно или външно – в дадена част от тялото. **Експериенцерът** в българския език, от своя страна, може да бъде кодиран на изреченско ниво като акузатив (36.), датив (37.) и притежание (38.) (Кукова / Кикова 2022: 45). Реализацията на **Експериенцера** допуска и удвояване, което подсилва изказването или разграничава говорещия от слушателя (39.). В редки случаи **Експериенцерът** е изразен със субект (40.) (Кукова / Кикова 2022: 46). Усещането може да се възприема като състоящо се в обобщена част от тялото (40) или да се конкретизира с предложна фраза (41.).

36. *Вратът все още [го]_{Експериенцер} БОЛЕСЕ.*

37. *Ножът ми беше под мен, дръжката [ми]_{Експериенцер} УБИВАШЕ на бедрото.*

38. *Защото напоследък ревматизмът [ми]_{Експериенцер} СЕ ОБАЖДА и усетих, че остарявам.*

39. *От нашия зъбобол [ни]_{Експериенцер} БОЛИ [нас и само нас]_{Експериенцер}.*

40. *[Танис]_{Експериенцер} ИЗПИТВАШЕ БОЛКА при всяко движение на тялото си.*

41. *При всички случаи ИМАХ БОЛКИ [в коремната област]_{Част на тялото} и се чувствах замаяна.*

В редки случаи се наблюдава известна вариативност в експлицитната проекция на **Експериенцера**, обикновено когато е известен и ясен за събеседниците (42. – 43.).

42. *Къде те БОЛИ?*

43. *Къде БОЛИ?*

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози |
|--------------------------------|---|---|--------------------------|------------------------|-----------------------|
| Експериенцер | Одушевената същност, която изпитва физическо усещане – вътрешно или външно – в дадена част от тялото. | eng-30-00007846-n: {човек}; eng-30-00015388-n: {животно} | NP | Подлог, Допълнение | – |
| Част на тялото | Мястото върху тялото, където се осъществява физическото усещане. | eng-30-05299178-n: {сетивен орган} | NP, PP | Подлог | – |
| | | | | Допълнение | В, под, на |

Таблица 4. Ядрени фреймови елементи на семантичния фрейм **Перцепция посредством тялото**

4.5. Създаване на шум⁵

Дефиниция във Фреймнет: „Физическа същност, наречена **Източник на звука (Sound source)** издава **Звук (Sound)**“. Това включва хора и животни, които издават звук с гласовите си органи. Този семантичен фрейм обхваща многобройни и разнородни глаголни лексикални единици. Свързан е с фрейма **Перцепция** чрез релацията **Използва**. Характерни за него са стимул-субектни глаголи за слухово възприятие, а именно:

- eng-30-02183175-v: {бибиткам; свиря};
- eng-30-02183787-v: {прокънтявам; прокънтя; ехтя; еча};
- eng-30-02172888-v: {шумя; про шумявам; про шумя};
- eng-30-02185861-v: {тракам; потраквам; потракам};
- eng-30-02173336-v: {свистя};
- eng-30-02187171-v: {пиукам};
- eng-30-02173838-v: {дрънкам; дрънвам; дрънна};
- eng-30-02185988-v: {гърмя; пукам};
- eng-30-02198014-v: {тътна; бумтя}.

⁵ Семантичният фрейм не включва ситуации, при които звукоизвличането е в резултат от контакт на предмети (сблъсък или триене). Случаите, в които одушевен **Агент** или природна сила причиняват звукоизвличането, се разглеждат от семантичния фрейм **Предизвикване на шум** – вж. 4.7. Ситуациите, в които звукът е свързан с движението на **Темата**, са разгледани в семантичния фрейм **Шум от движение** – вж. 4.6.

Ядрените фреймови елементи **Шумно събитие (Noisy event)**, **Източник на звук** и **Звук** са дефинирани във Фреймнет, както следва:

Шумно събитие е фреймовият елемент, назоваващ събитие, в което участват един или повече **Източници на звук** и е придружено от **Звук** (44. – 45.).

44. [*Смехът на гостите*]_{Шумно събитие} *КЪНТЯ* дълго.

45. [*Стъпките им*]_{Шумно събитие} *ОТЕКНАХА* в тъмното.

Звук – фреймовият елемент много често е инкорпориран в значението на глаголите, на които е приписан фреймът. Понякога се среща като именна фраза, изразена със *звук* или подредните му. **Звукът** може да бъде кодиран и посредством субекта на изречението (46. – 47.).

46. *В ушите ѝ ЕХТЕШЕ* [*силен тътен*]_{Звук}

47. *ГЪРМЯХА* [*гръмотевици*]_{Звук} и *светкавици* *разбиваха вражеските строежи с копията си.*

Източник на звука – назовава физическата същност, която издава звук. Този фреймов елемент включва както одушевени, така и неодушевени физически същности, задължително със сложна вътрешна структура, които могат да издават звук (човек, животно, машина, музикален инструмент и др.) (48.).

48. *В кухнята ШУМЕШЕ* [*телевизорът*]_{Източник на звука} и се чуваше *приповдигнатият тон на мъжа, който съобщаваше прогнозата за времето и напомняше на зрителите, че Чарлстън все още е в разгара на топлинна вълна, каквато не е сполитала града от трийсет години.*

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози (свързващи думи) |
|--------------------------|--|---|-----------------------|---------------------|-------------------------------------|
| Шумно събитие | Фреймовият елемент, назоваващ събитие, в което участват един или повече Източници на звук и е придружено от Звук . | eng-30-07371293-p: {звук}; eng-30-07387509-p: {шум}; eng-30-07020895-p: {музика}; eng-30-07109196-p: {реч} | NP | Подлог | — |

| | | | | | |
|--------------------------|---|--|----|--------|---|
| Звук | Звуковият сигнал, излъчван от Източника или придружаващ Шумното събитие . | eng-30-05718254-n: {звук} | NP | Подлог | — |
| Източник на звука | Фреймов елемент, назоваващ физическата същност, която издава Звук . | eng-30-00007846-n: {човек}; eng-30-00015388-n: {животно}; eng-30-03575240-n: {инструмент} | NP | Подлог | — |

Таблица 5. Ядрени фреймови елементи на семантичния фрейм **Създаване на шум**

4.6. Шум от движение

Дефиниция във Фреймнет: Този семантичен фрейм обхваща глаголи, „назоваващи звук, който характеризира движение“.

Свързан е с фреймовете **Движение (Motion)** и **Създаване на шум** с помощта съответно на релациите **Наследяване** и **Виж също**. Типични представители са:

- eng-30-02184797-v: {стържа; скрипя};
- eng-30-02182109-v: {бръмча; жуца};
- eng-30-02179279-v: {жвакам};
- eng-30-02178709-v: {издрънчавам; издрънча};
- eng-30-02179372-v: {пухтя};
- eng-30-02187922-v: {ромоля; ромоня; бълбукам};
- eng-30-02184965-v: {трополя}.

Ядрени фреймови елементи в рамките на този фрейм са **Тема (Theme)**, **Област (Area)**, **Цел (Goal)**, **Път (Path)**, **Начална точка (Source)** и са дефинирани, както следва:

Тема – назовава онази физическа същност, която променя местоположението си (49.). Може да бъде външен аргумент или ако има друга причина за движението – NP обект.

49. [Цели облаци комари]_{Тема} **ЖУЖАХА** в гьсталака от другата страна на дерето. Бръмчаха. И виеха.

Област – назовава мястото на извършване на движението, особено ако то се възприема като следващо сложна, нелинейна траектория. Фреймовият елемент **Област** най-често се въвежда от предлога *из* (50.).

50. *Някакъв едър бръмбар доста дълго БРЪМЧА [из стаичката му]_{Област} докато най-сетне слепешком улучи прозореца.*

Цел – фреймовият елемент посочва къде се озовава **Темата** в резултат от движението. Въвежда се обичайно с предлозите *на* или *върху* (51. – 52.).

51. *Брадвата ИЗДРЪНЧА [на каменните плочи]_{Цел} и отекна като камбана във внезапната тишина.*

52. *Нещо тежко ИЗТРОПОЛЯ [върху покрива]_{Цел}; всички свеци по первазите подскочиха.*

Двата фреймови елемента **Област** и **Цел** са взаимноизключващи се компоненти и не се срещат в рамките на едно и също изречение.

Път – посочва траекторията на движението, произвеждащо звук (53.). Различава се от **Цел** или **Начална точка**.

53. *Нямам нерви да чакам асансьора, затова тръгвам по стълбите и ТРОПОЛЯ [надолу по стъпалата]_{Път}.*

Начална точка – назовава началната точка, откъдето започва движението, предизвикващо звука (54.). **Началната точка** в този фрейм е с много рядка употреба в българския език.

54. *Товарните коли ИЗТРОПОЛЯХА [извън портите на хана]_{Начална точка}.*

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози (свързващи думи) |
|--------------------------|--|--|-----------------------|---------------------------|-------------------------------------|
| Област | Мястото на извършване на движението, особено ако то се възприема като следващо сложна, нелинейна траектория. | eng-30-02735688-n: {пространство}; eng-30-08664443-n: {място} | PP | Обстоятелствено пояснение | Из, в, във |
| Цел | Фреймовият елемент посочва къде се озовава Темата в резултат от движението. | eng-30-00002684-n: {материално тяло} | PP | Обстоятелствено пояснение | Върху, на |

| | | | | | |
|----------------------|--|--|----|-----------------------|-----------|
| Път | Последователност от точки в пространството, която посочва траекторията на движението, произвеждащо Звук . | eng-30-02735688-n: {пространство}; eng-30-08664443-n: {място} | PP | Обстоятелно пояснение | През, по |
| Начална точка | Пространствена точка, която назовава мястото, откъдето започва движението, предизвикващо Звук . | eng-30-00002684-n: {материално тяло} | PP | Обстоятелно пояснение | От, извън |
| Тема | Физическа или абстрактна същност, която променя местоположението си, като издава звуков сигнал. | eng-30-00015388-n: {животно}; eng-30-00021939-n: {артефакт} | NP | Подлог, допълнение | |

Таблица 6. Ядрени фреймови елементи на семантичния фрейм **Шум от движение**

4.7. Предизвикване на шум

Дефиниция във Фреймнет: „Агент (Agent) или Причина (Cause) каузират издаване на звук от даден **Издаващ звук (Sound maker)**“. Семантичният фрейм е свързан с релацията **Наследяване** с фрейма **Транзитивно действие (Transitive action)** и с релацията **Каузация** с фрейма **Създаване на шум**. Типични глаголи, проектиращи този фрейм, са:

eng-30-02186690-v: {чуквам; чукна};
eng-30-02181538-v: {звъня; бия};
eng-30-02135727-v: {надувам; надуя}.

Ядрените фреймови елементи са дефинирани, както следва:

Агент – индивидът, който манипулира определен **Издаващ звук**, така че да произведе **Звук** (55).

Издаващ звук – физическата същност, която произвежда звук в резултат на дейността на **Агент** или **Причина** (55).

55. *Един час преди полунощ [Ханс]_{Агент} тръгна да БИЕ [камбаната]_{Издаващ звук}.*

Причина – неодушевена природна същност, която прави така, че **Издаващият звук** да произведе звук (56.).

56. [Такситата]_{Причина} **НАДУВАХА** [клаксоните]_{Издаващ звук} и се стрелкаха между по-бавните коли и камионите.

И трите фреймови елемента са ядрени, като **Причината** рядко се появява в едно изречение заедно с неядрения елемент **Средство (Means)**.

| Ядрени фреймови елементи | Дефиниция | Семантични ограничения | Синтактична категория | Синтактична функция | Допустими предлози (свързващи думи) |
|--------------------------|--|---|-----------------------|---------------------|-------------------------------------|
| Агент | Индивидът, който манипулира определен Издаващ звук , така че да произведе Звук . | eng-30-00007846-n: {човек}; eng-30-00015388-n: {животно} | NP | Подлог | — |
| Причина | Неодушевена природна същност, която прави така, че Издаващият звук да произведе Звук . | eng-30-00002684-n: {материално тяло}; eng-30-07283608-n: {явление} | NP | Подлог | — |
| Издаващ звук | Физическата същност, която произвежда Звук в резултат на дейността на Агент или Причина . | eng-30-00021939-n: {артефакт} | NP, PP | Допълнение | По, на, с, със |

Таблица 7. Ядрени фреймови елементи на семантичния фрейм **Предизвикване на шум**

5. Заключение и бъдеща работа

Изследването е посветено на основни глаголи за перцепция и семантичните фреймове от Фреймнет, в които се реализират. Подробно са анализирани фреймовете им елементи и са описани семантичните, синтактичните и морфологичните им параметри. Раздел 2. представлява описание на използваните ресурси със съдържателните и структурните им характеристики, като очертава понятието за концептуален фрейм и компонентите му. Раздел 3. обхваща авторитетни досегашни изследвания и мотивацията, която оформя проучването. Представени са различни подходи към класификацията на предикатите за перцепция, които служат за основа на настоящото изследване. Систематизирани са възможни бъдещи приложения на обобщенията от проучването. В Раздел 4. е съсредоточен основният аналитичен преглед на семантичните фреймове, в които се реализират перцептивните предикати. Детайлно са представени седем фрейма – **Перцепция**, **Активна перцепция**, **Пасивна перцепция**, **Перцепция посредством тялото**, **Създаване на шум**, **Шум от движение** и **Предизвикване на шум**. Прототипният фрейм **Перцепция** служи като основа за описание на останалите семантични фреймове, които са свързани чрез различни релации с него, тъй като въвежда основните участници в перцептивната ситуация – **Възприемащ** и **Феномен**. Следващите два фрейма (**Активна перцепция** и **Пасивна перцепция**) обобщават синтактико-семантичните реализации на експериментално-субектните глаголи съответно за волева и пасивна перцепция. **Предизвикване на шум** е обвързан с глаголите за перцепция посредством многозначността им и е включен в прегледа, тъй като онагледява паралелните реализации на различни значения на едни и същи по форма глаголи. Семантичните фреймове **Перцепция посредством тялото**, **Създаване на шум** и **Шум от движение** демонстрират синтактичното поведение на стимул-субектните глаголи за перцепция.

Резултатите от изследването показват, че различните проявления на фреймовете елементи (синтактични и семантични) са отражение на семантиката на перцептивните събития и на ясната им езикова категоризация. Тази системна закономерност е отразена по недвусмислен начин в семантичните и граматическите отношения между глаголни лексикални единици и конкретните прояви на фреймовете им елементи, като им придава солидна семантична мотивация.

Благодарности: Изследването е осъществено по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания – 2020 г.“, Договор No. КП-06-Н50/1 от 2020 г.

Литература / References

- Вежбицкая 1986: Вежбицкая А. Восприятие: семантика абстрактного словаря – В: *Новое в зарубежной лингвистике*, вып. 18, с. 336 – 370. (Vezhbickaya 1986: Vezhbickaya A. Vospriyatie: semantika abstraktnogo slovarya. – In: *Novoe v zarubezhnoj lingvistike*, vol. 18, pp. 336 – 370.)
- Коева 2010: Коева, Св. *Българският Фреймнет*. София: Архимед, 146 с. (Коева 2010: Koeva, Sv. *Balgarskiyat Framenet*. Sofia: Arhimed, p. 146.)
- Кукова 2022: Кукова, Х. *Формално описание на перцептивните и менталните предикати в българския език (снемане на семантична многозначност)* (автореферат). <https://ibl.bas.bg/wp-content/uploads/2022/05/H_Kukova_Avtoreferat_BG.pdf> [25.03.2024] (Kukova 2022: Kukova, H. *Formalno opisanie na pertseptivnite i mentalnite predikati v balgarskiya ezik (snemane na semantichna mnogoznachnost)* (avtoreferat). <https://ibl.bas.bg/wp-content/uploads/2022/05/H_Kukova_Avtoreferat_BG.pdf> [25.03.2024].)
- Лесева и др. 2018: Лесева, Св., Ив. Стоянова, Хр. Кукова, М. Тодорова. Интегриране на субкатегоризационна информация в реляционната структура на УърдНет. – *Български език*, 65, № 2, с. 13 – 40. (Leseva et al. 2018: Leseva, S., I. Stoyanova, H. Kukova, M. Todorova. Integrirane na subkategorizatsionna informatsiya v relationsonnata struktura na WordNet. – *Balgarski ezik*, 65, № 2, pp. 13 – 40.)
- Моисеева 2005: Моисеева, С. А. *Семантическое поле глаголов восприятия в западно-романских языках*. Белгород: Изд-во БелГУ, с. 248. (Moiseeva 2005: Moiseeva, S. A. *Semanticheskoe pole glagolov vospriyatiya v zapadno-romanskih yazykah*. Belgorod: BelGU, p. 248)
- Стефанов 2024: Стефанов, С. *Опит за обяснение на една глаголна употреба в разговорния български език през призмата на етимологията* (под печат). (Stefanov 2024: Stefanov, S. *Opit za obyasnienie na edna glagolna upotreba v razgovorniya balgarski ezik prez prizmata na etimologiyata* (pod pechat).)
- Baker et al. 1998: Baker, C. F., C. Fillmore, J. Lowe. The Berkeley FrameNet Project. – In: *Proceedings of COLINGACL*. Montreal, Canada: Association for Computational Linguistics, pp. 86 – 90.
- Baker et al. 2003: Baker, C. F., C. Fillmore, B. Cronin. The Structure of the FrameNet Database. – *International Journal of Lexicography*, vol. 16, pp. 281 – 296.
- Fellbaum 1998: Fellbaum, C. (Ed.) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.
- Fillmore 2002: Linking Sense to Syntax in FrameNet. – In: *Proceedings of 19th International Conference on Computational Linguistics*. Taipei: Association for Computational Linguistics.
- Koeva 2020: Koeva, S. Semantic Relations and Conceptual Frames. – In: Koeva, S. (Ed.) *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia: Prof. Marin Drinov Academic Publishing House of the Bulgarian Academy of Sciences, pp. 7 – 20.
- Koeva 2021a: Koeva S. The Bulgarian WordNet: Structure and specific features. – *Papers of BAS. Humanities and Social Sciences*, vol. 8, no. 1, pp. 47 – 70.
- Koeva 2021b: Koeva S. Towards Expanding WordNet with Conceptual Frames. – In: *Proceedings of the 11th Global Wordnet Conference, University of South Africa (UNISA)*. Global Wordnet Association, pp. 182 – 191.

- Koeva, Doychev 2022: Koeva, S., E. Doychev. Ontology Supported Frame Classification. – In: *Proceedings of the 5th International Conference on Computational Linguistics in Bulgaria (CLIB 2022)*, pp. 203 – 231.
- Koeva et al. 2012: Koeva, S., I. Stoyanova, S. Leseva, Ts. Dimitrova, R. Dekova, and E. Tarpomanova. The Bulgarian National Corpus: Theory and Practice in Corpus Design. – *Journal of Language Modelling*, no. 1, pp. 65 – 110.
- Leseva et al. 2019: Leseva, S., I. Stoyanova, M. Todorova, H. Kukova. A Theoretical Overview of Conceptual Frames and Semantic Restrictions on Frame Elements. – *Linguistique balkanique*, no. 2, pp. 172 – 187.
- Levin 1993: Levin, B. *English verb classes and alternations: A preliminary investigation*. Chicago: The University of Chicago Press.
- Miller 1995: Miller, G. A. WordNet: A lexical database for English. – *Communications of the ACM*, vol. 38, no. 11, pp. 39 – 41.
- Ruppenhofer et al. 2016: Ruppenhofer, J., M. Ellsworth, M. R. L. Petruck, C. R. Johnson, C. F. Baker, J. Scheffczyk. *FrameNet II: extended theory and practice*. Berkeley, California: International Computer Science Institute.
- Verhoeven 2007: Verhoeven, E. *Experiential constructions in Yucatec Maya: a typologically based analysis of a functional domain in a Mayan language*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Viberg 1983: Viberg A. The Verbs of Perception: A Typological Study. – *Linguistics*, vol. 21, no. 1, pp. 123 – 162.
- Viberg 2001: Viberg, A. Verbs of perception. – In: M. Haspelmath, E. König, W. Oesterreicher, W. Raible (Eds.). *Language Typology and Language Universals: An International Handbook*. Berlin: W. de Gruyter, pp. 1294 – 1309.

MAIN PERCEPTION FRAMES AND COMPLEX PRESENTATION OF SEMANTIC AND SYNTACTIC INFORMATION

Hristina Kukova

Institute for Bulgarian Language “Prof. Lyubomir Andreychin”

hristina@dcl.bas.bg

Abstract. The research focuses on verbs of perception and their conceptual descriptions in Bulgarian. An analytical review of previous theoretical generalisations as well as scientifically applied research is presented. The variety of constructions and syntactic realisations that the representatives of this class of predicates project onto their sentential surroundings serves as a starting point for the present study. The analysis is based on semantic and partly morphological information about the lexical units from WordNet (Fellbaum 1998; Koeva 2021a), as well as on semantic-syntactic specificities by which the lexical units under study are represented in FrameNet (Baker et al. 1998; Ruppenhofer et al. 2016). Seven frames are considered, which demonstrate high frequency and a wide

variety of verbal lexical units. The description includes grammatical features of the verbs, semantic and syntactic constraints that the verbal lexical units impose on their frame elements, mapping of the frame elements to a certain synonym set or semantic class of WordNet nouns that corresponds to the properties of the elements in the frame. The status of the frame elements is also adopted from FrameNet, which is essential for the realisation of the given lexical unit.

Keywords: *perception verbs; semantic frames; frame elements*

Hristina Kukova
Institute for Bulgarian Language
Bulgarian Academy of Sciences
52 Shipchenski prohod Blvd., bl. 17
Sofia 1113, Bulgaria
<https://orcid.org/0000-0001-9938-546>

СЕМАНТИЧНИ СТУДИИ.
СИНТАКТИЧНО И СЕМАНТИЧНО ОПИСАНИЕ НА БЪЛГАРСКИ ГЛАГОЛИ

БЪЛГАРСКА

Първо издание

Редактор: Светла Коева
Коректор: Мила Микова
Художник на корицата: Божидар Чимширов
Графичен дизайнер: Вероника Томчева

Формат 70×100/16

Печатни коли 11,25

Печатница на Издателство на БАН „Проф. Марин Дринов“
1113 София, ул. „Акад. Г. Бончев“, бл. 5

www.press.bas.bg <http://ibl.bas.bg/>

ISBN 78-619-245-448-7
ISBN online 78-619-245-448-7

