

## **Семантико-синтактичен интерфейс: универсалност и специфичност**

### **СЕМАНТИЧНИ И КОНЦЕПТУАЛНИ ФРЕЙМОВЕ: УНИВЕРСАЛНИ СВОЙСТВА И СПЕЦИФИЧНИ ЧЕРТИ**

*СВЕТЛА КОЕВА (СОФИЯ, БЪЛГАРИЯ)*

**Abstract.** The study's goal is to show how semantic frame and conceptual frame, two abstract semantic structures used to describe language, are interrelated. More specifically, it demonstrates how much of the structure of the semantic frames is inherited in the conceptual frames, as well as the universal and specific aspects of the two theoretical frameworks.

In this regard, the semantic frames of FrameNet and the conceptual frames used to develop Bul-Frame (the Bulgarian version of FrameNet) are briefly described. The four correspondence models between semantic and conceptual frames are discussed: equivalence, inheritance and equivalence, partial equivalence, and no equivalence. Conceptual frames linked by one of the three types of equivalence are organised into a superframe that, at an abstract level, corresponds to the respective semantic frame.

The study demonstrates when language-specific conceptual frames for Bulgarian are required, as well as which elements of the conceptual frames reflect characteristics unique to Bulgarian verbs. Information about the set of nouns that are suitable for language realisation of the frame elements is a distinguishing property of conceptual frames.

The research is aimed at describing the semantic and syntactic structure of Bulgarian verbs in a format comparable to the description of semantic frames in English and other languages. This enables the integration of the semantic representation of the Bulgarian language in a global network, displaying the unique semantic and syntactic characteristics of individual languages as well as universal properties that are valid for all or a subset of the described languages.

**Keywords:** semantic frame, conceptual frame, FrameNet, interlanguage equivalence

#### **I. ВЪВЕДЕНИЕ**

Целта на настоящото изследване е да представи как се съотнасят две абстрактни семантични структури, които се използват за описание на езика: семантичен фрейм и концептуален фрейм. По-точно, ще бъде показано каква част от структурата на семантичните фреймове се наследява в концептуалните фреймове и кои са универсалните свойства и специфичните характеристики на двете теоретични рамки.

Във връзка с това накратко ще бъдат представени семантичните фреймове на Фреймнет<sup>1</sup> (Johnson, Fillmore 2000; Fillmore, Baker 2010; Ruppenhofer et al.,

---

<sup>1</sup> <https://framenet.icsi.berkeley.edu>

2016) и концептуалните фреймове (Коева 2020, 2021; Коева, Douchev 2022), използвани за целите на изграждането на Булфрейм<sup>2</sup> (българското съответствие на Фреймнет). Ще бъдат описани четирите модела за съответствие между семантичен и концептуален фрейм: еквивалентност, наследяване и еквивалентност, частична еквивалентност и липса на еквивалентност. Концептуалните фреймове, които са свързани с една от трите форми на еквивалентност, са организирани в суперфрейм, който на абстрактно равнище съответства на даден семантичен фрейм.

Изследването е насочено към описанието на семантичната и синтактичната структура на българските глаголи във формат, съотнесим с описанието на семантичните фреймове за английски и други езици. Това позволява включването на семантичното представяне за български език в глобална мрежа, показваща уникалните семантични и синтактични характеристики на отделните езици, както и универсалните свойства, които са валидни за всички или за група от описаните езици.

## II. СЕМАНТИЧЕН ФРЕЙМ

Фреймнет организира лексикалните единици (двойката от дума или съставна дума и нейното значение), като ги свързва с абстрактно семантично представяне – семантичен фрейм. Семантичният фрейм се дефинира като концептуална структура, която описва определен тип ситуация, обект или събитие, заедно с неговите участници и свойства (Johnson, Fillmore 2000: 56; Ruppenhofer et al. 2016: 7). Приема се, че значението на думите се проявява в контекста на фреймовете, с които те се асоциират (Fillmore 1968, 1977, 1982). Различните значения на думите са свързани с различни семантични фреймове и дадено значение се проявява само в рамките на един фрейм.

Семантичните фреймове се характеризират с име на фрейма, дефиниция, семантичен тип, релации между фреймовете и съдържат елементи на фрейма, които също имат име, дефиниция, семантичен тип, спецификация за техния ядрен статус и релации помежду си. Например семантичният фрейм **Cutting**<sup>3</sup> (**Рязане**) се описва с дефиницията: *An **Agent** cuts an **Item** into **Pieces** using an **Instrument** (which may or may not be expressed) (**Агент** реже (цял) **Обект** на **Части**, като използва **Инструмент** (който може да не бъде изразен)). Лексикалните единици, които характеризират фрейма в английски са: *carve, chop, cube, cut, fillet, mince, slice* (изрязвам, нарязвам, нарязвам на кубчета<sup>4</sup>, режа, филетирам, кълцам, нарязвам на тънки ивици<sup>5</sup>). Ядрените фреймови елементи са: **Agent**: *The Agent is the person cutting the Item into Pieces* (**Агент**: Лицето, разделящо (целия) **Обект** на части); **Item**: *The item which is being cut**

<sup>2</sup> <https://dcl.bas.bg/bulframe/>

<sup>3</sup> <https://framenet.icsi.berkeley.edu/frameIndex>

<sup>4</sup> Няма лексикализация на български.

<sup>5</sup> Няма лексикализация на български.

into Pieces (**Обект: Обектът**, който се разделя на **Части**); **Pieces: The Pieces are the parts of the original Item which are the result of the slicing** (**Части: Частите** от първоначалния **Обект**, които се получават в резултат на рязането).

В допълнение, между семантичните фреймове са формулирани релации, а за реализацията на фреймовите елементи е предвидена синтактична анотация. Например фреймът **Рязане** е свързан с релацията **наследяване** с фрейма **Intentionally\_affect** (**Съзнателно въздействие**) с дефиниция *An Agent causes a Patient to be affected, sometimes by a particular Means or by use of an Instrument* (**Агент** прави така, че **Търпящ** да бъде засегнат, понякога с определени **Средства** или като използва **Инструмент**). Фреймнет може да се разглежда като семантична мрежа (или набор от малки семантични мрежи), чиито възли съставляват семантичните фреймове и чиито дъги изразяват семантични отношения между тях (Коева 2020: 8).

Фреймовите елементи са разделени на ядрени, периферни и екстратемотични. Ядреният елемент съответства на концептуално задължителен компонент на фрейма, като същевременно го прави уникален и различен от останалите фреймове (Ruppenhofer et al. 2016: 23–24). Ядрените елементи на фрейма може да са взаимосвързани по различен начин (Ruppenhofer et al. 2016: 25): присъствието на който и да е елемент от дадено множество е достатъчно, за да се изрази значението на предиката, възможна е и съвместната им употреба (*пътувам [от Пловдив][за София]*); появата на даден елемент изисква задължителна поява на друг (*\*завързах [за стола]; завързах [крадеца] [за стола]*); появата на елемент от дадено множество изключва появата на останалите елементи (*измъкнах се [от преследвачите]; измъкнах се [от ареста]*).

Във Фреймнет се дефинират семантични типове на елементите на фрейма (Ruppenhofer et al. 2016: 12), например [Съзнателно същество], [Физически обект] и под. Не всички елементи са снабдени със семантичен тип и семантичните типове в много случаи не показват действителните ограничения за лексикална съчетаемост.

### III. КОНЦЕПТУАЛЕН ФРЕЙМ

Концептуалният фрейм също може да се дефинира като абстрактна структура, която описва определен тип ситуация или събитие, заедно с неговите участници и свойства (Коева, 2020: 7). Основната разлика между семантичен и концептуален фрейм се изразява в това, че елементите на концептуалния фрейм се асоциират с множеството от лексикални единици, посредством които потенциално могат да се реализират. В този смисъл и семантичният фрейм може да се нарече концептуален, както и обратно, но терминът **концептуален фрейм** се въвежда, за да се разграничат двете теоретични рамки.

Концептуалните фреймове за български (към момента) се формулират само за глаголи. Те се характеризират с име на фрейма, дефиниция, семантичен тип, релации между фреймовете; съдържат елементи на фрейма, които,

от своя страна, също имат име, дефиниция, семантичен тип, ядрен статус и релации помежду си. Тази информация се наследява от семантичните фреймове, ако такива вече са дефинирани и ако информацията е валидна за българския език.

Глаголите, които се описват в рамките на даден концептуален фрейм, могат да бъдат един или няколко, свързани помежду си със семантични отношения на лексикално равнище (синонимия, антонимия) и/или йерархични семантични отношения на концептуално равнище (хиперонимия, тропонимия, каузация). Например глаголите *режа*, *разрежа*, *разрязвам* са включени в едно синонимно множество в Българския уърднет<sup>6</sup> със значение 'разделям нещо на части с остър инструмент'. Хиперонимно синонимно множество образуват глаголите *разделям*, *разделя* със значение 'правя нещо на части, дялове', а сред хипонимните синонимни множества са глаголите *нарежа*, *нарязвам* със значение 'режа нещо на по-малки части, парчета'; *сека*, *отсека*, *отсичам* със значение 'режа нещо със замах с остър инструмент, за да го разделя на части, отделя' и др. Всички изброени глаголи се асоциират с концептуалния фрейм **Рязане**. При описанието всеки глагол получава морфологична спецификация в зависимост от категориите лице и число; вид на глагола и присъединяването на пряк обект (транзитивност).

Елементите на концептуалния фрейм до голяма степен съответстват на ядрените елементи на семантичните фреймове от Фреймнет (ако това е приложимо за български). Вместо синтактична анотация се включва синтактично описание на реализацията на фреймовите елементи. Например *режа* от фрейма **Рязане** се характеризира със следната синтактична анотация на фреймовите елементи:

**Агент:** NP (именна група), субект, свободно изпускане

**Обект:** NP (именна група), обект, незадължителна реализация

**Части:** PP (предложна група), обект (*на*), незадължителна реализация

**Инструмент:** PP (предложна група), обект (*с*), незадължителна реализация

Всеки ядрен елемент на концептуалния фрейм е съотнесен с множество от съществителни, които са съвместими за употреба с глаголите, за които се отнася фреймът. Множеството може да съдържа само едно съществително, няколко или голям брой съществителни, свързани помежду си със семантични отношения на лексикално равнище (синонимия, антонимия) или с йерархични концептуални отношения (хиперонимия, хипонимия). Например глаголът *варя* със значение 'приготвям храна или под. в много гореща или вряща вода' предизвиква представата за фрейма **Apply\_heat (Топлинна обработка)**, който се характеризира с четири фреймови елемента **Готвач**, **Храна**, **Контейнер** и **Инструмент за топлинна обработка**, като за всеки един от тях е специфици-

<sup>6</sup> <https://dcl.bas.bg/bulnet/>

рано синонимното множество (едно или повече) от Българския уърднет, което доминира над подходящи за съвместна употреба съществителни имена.

**Готвач:** NP (именна група), субект, свободно изпускане, *човек* (eng-30-00007846-n)

**Храна:** NP (именна група), обект, незадължителна реализация, *храна* (eng-30-07555863-n); *месо* (eng-30-07649854-n); *риба* (eng-30-07775375-n); *зеленчук* (eng-30-07707451-n)

**Контейнер:** PP (предложна група), обект (*в*), незадължителна реализация, *тенджер* (eng-30-03990474-n)

**Инструмент за топлинна обработка:** PP (предложна група), обект (*на, върху*), незадължителна реализация, *огън* (ng-30-03343560-n); *котлон* (eng-30-03543254-n); *огнище* (eng-30-08581699-n)

В заключение, концептуалните фреймове се отличават с директно морфологично и синтактично описание на елементите на фрейма (без анотация на примери) и с информация за семантичната съчетаемост между глаголите и съществителните, които могат да се реализират като съответните фреймови елементи.

#### IV. НАСЛЕДЯВАНЕ НА ЕЗИКОВО НЕЗАВИСИМА ИНФОРМАЦИЯ ОТ СЕМАНТИЧНИТЕ ФРЕЙМОВЕ

При наследяването на езиково независимата информация, дефинирана в семантичните фреймове, може да съществуват следните четири случая по отношение на концептуалните фреймове: еквивалентност, наследяване и еквивалентност, частична еквивалентност, липса на еквивалентност. Наследяването се отнася само за информацията, която е налична във Фреймнет и може или не може да бъде наследена при семантично описание на българските глаголи.

Релация на междуезикова еквивалентност се наблюдава, когато в двата езика съществуват лексикални единици (преводни еквиваленти), които се съотнасят със семантичната информация, дефинирана в семантичен фрейм и пренесена в концептуалния фрейм. Например семантичният фрейм **Breaking\_apart (Разделяне на части)** с дефиниция *A Whole breaks apart into Pieces, resulting in the loss of the Whole (and in most cases, no piece that has a separate function)* (**Цяло** се разделя на **Части**, в резултат на което престава да съществува като **Цяло**) се отнася както за английските лексикални единици *break apart, break, crumble, fragment, shatter, snap, splinter*, така и за българските им преводни еквиваленти *чупи се, счупи се, счупва се, разпада се, разпадне се, разтрошава се, разтроши се, строшава се, строши се*. Синонимите от синонимното множество *счупи се, счупва се* със значение 'за твърдо тяло или предмет – разделям се, разпадам се на части или парчета' (както и други синонимни множества) наследяват семантичния фрейм от Фреймнет и съответните фреймови елементи:

**Цяло:** NP (именна група), субект, свободно изпускане, *артефакт* (eng-30-00021939-n)

**Части:** PP (предложна група), обект (*на*), незадължителна реализация, *артефакт* (eng-30-00021939-n); *част* (eng-30-09385911-n)

**Наследяване и релация на междуезикова еквивалентност** се наблюдава, когато във Фреймнет е дефиниран семантичен фрейм, който е подходящ за описанието на български глагол, но във Фреймнет липсва негов преводен еквивалент. Например семантичният фрейм **Разделяне на части** е приложим и за описание на синонимното множество *спуква се, спуска се* със значение 'за твърдо тяло или предмет – разделя се, разпада се внезапно и със сила на части', което е хипоним на синонимното множество *счупи се, счупва се* и релацията между двете синонимни множества позволява да се наследи семантичният фрейм на хиперонима.

**Релация на частична междуезикова еквивалентност** се наблюдава, когато във Фреймнет е дефиниран семантичен фрейм, който частично е подходящ за описанието на български глагол. В такъв случай може да има изключване, замяна или допълване на фреймов елемент. Например семантичният фрейм **Разделяне на части** е подходящ частично за описанието на синонимите *пробива се, пробие се* със значение 'за твърдо тяло или предмет – претърпява нарушаване на целостта чрез намушкване, продупчване с остър предмет', при което се реализира само един фреймов елемент:

**Цяло:** NP (именна група), субект, свободно изпускане, *артефакт* (eng-30-00021939-n)

За да се запази връзката със семантичния фрейм, който се реализира при изразяването на значението на български глагол посредством (наследяване на) междуезикова еквивалентност и частична междуезикова еквивалентност, както и да се отрази връзката между семантично свързани глаголи, които имат различни морфологични и синтактични характеристики или различна съчетаемост със съществителни имена, се въвежда абстрактната семантична структура (**концептуален**) **суперфрейм**. Концептуалният суперфрейм е еквивалентен на (езиково независимата информация от) семантичния фрейм от Фреймнет и неговата роля е да обединява концептуалните фреймове, които напълно или частично споделят семантичната информация от Фреймнет. Концептуалните фреймове, които се съотнасят с даден суперфрейм, се характеризират с името на семантичния фрейм и допълнително уникално име, съвпадащо с един от глаголите, за които се отнася фреймът. Например езиково независимата информация от семантичния фрейм **Breaking apart** се пренася в суперфрейма **Разделяне на части**, който от своя страна се свързва с концептуалните фреймове **Разделяне на части** (*чупи се*) и **Разделяне на части** (*пробива се*), които напълно или частично наследяват тази езиково независима информация.

**Липса на междуезикова еквивалентност** се наблюдава, когато във Фреймнет не е дефиниран семантичен фрейм, подходящ за описанието на български глагол, и се дефинира нов. Това може да се случи по две причини:

Понятието е концептуализирано в английски, но все още не е създаден съответният семантичен фрейм във Фреймнет (лъжлива липса), например *to golf* (играя голф).

Понятието не е концептуализирано в английски (същинска липса), например *захъркам, захърквам, затърся, затърсвам*.

## **V. СРАВНЕНИЕ МЕЖДУ СЕМАНТИЧЕН И КОНЦЕПТУАЛЕН ФРЕЙМ**

Езиково независимата информация в семантичните фреймове (наследена в концептуалните фреймове) е на семантично равнище: дефиницията на фрейма, релациите между фреймове, броят и видът на елементите на фрейма, както и тяхната дефиниция, семантичен тип, ядрен статус и релации. Наследява се и административна информация: името на фрейма и имената на фреймовите елементи. Наред с това концептуалните фреймове може да съдържат аналогична информация, която е езиково специфична, или да се характеризират с частична междуезикова еквивалентност. При всички случаи концептуалните фреймове съдържат информация, която се отнася до морфологичните и синтактичните характеристики на българските глаголи, както и семантична информация, която не се съдържа в семантичните фреймове – за множеството от съществителни имена, които са подходящи за реализация на отделните фреймови елементи. Съотношението при наследяването, уникалността и специфичността на информацията в семантичните и концептуалните фреймове схематично е представена в Таблица 1.

## **VI. ЗАКЛЮЧЕНИЕ**

В изследването се представят две съотносими теоретични рамки за представяне на семантична и синтактична информация: семантичните фреймове на Фреймнет, базирани на теорията за семантично описание на Чарлз Дж. Филмор, и концептуалните фреймове на Булфрейм – българския аналог на Фреймнет. Описани са случаите, при които валидната за двата езика семантична информация, кодирана в семантичните фреймове, се използва в структурата на концептуалните фреймове. Показва се кога се налага дефинирането на езиково специфични за български концептуални фреймове, както и кои части от концептуалните фреймове отразяват специфични за българските глаголи характеристики. Отличителна характеристика на концептуалните фреймове е наличието на информация за множеството от съществителни имена, които са подходящи за езикова реализация на фреймовите елементи.

Таблица 1. Сравнение между семантичен и концептуален фрейм

		Семантичен фрейм	Концептуален фрейм
Административна информация	Име на фрейма; име на елемент на фрейма	Уникална	Наследява се в суперфрейма или е уникална при езикова специфичност
	Идентификационен номер на глагола	Не	Идентификационният номер на синонимното множество от Уърднет
Лексикална информация	Информация за типа на глаголната лема	Не	дума, съставна дума
	Част на речта	Глаголи, съществителни, прилагателни, наречия	Глаголи
	Дефиниция	Уникална	Уникална, от Уърднет
	Стилистични бележки	Не	от Уърднет
	Релации с други глаголи	Не	от Уърднет
Грамматична информация	Парадигма на категориите лице и число за глаголите	Не	Личен, безличен, третоличен, множественоличен
	Съчетаемост с пряк обект за глаголите	Не	Транзитивен, интранзитивен
	Вид на глагола	Не	Свършен, несвършен, перфектива тантум, имперфектива тантум, двувидов
Семантична информация	Дефиниция на фрейма; релации на фрейма с други фреймове; брой на елементите на фрейма; статус, дефиниция и тип на елементите на фрейма; релации между елементите на фрейма	Езиково независима или езиково специфична за английски	Наследява се в суперфрейма или е езиково специфична
	Семантична съчетаемост при лексикална реализация на фреймовите елементи	Не	Синонимни множества от Уърднет, които доминират над множеството от съществителни, чрез които се реализират фреймовите елементи
Синтактична информация за лексикалната реализация на фреймовите елементи	Синтактична категория	Посредством анотация	Директно
	Синтактична функция	Посредством анотация	Директно
	Имплицитност	Посредством анотация	Директно



## БЛАГОДАРНОСТИ

*Разработката е осъществена по проекта „Обогатяване на семантичната мрежа Уърднет с концептуални фреймове“, подкрепен от Фонд „Научни изследвания“, Договор КП-Об-Н50/1 от 2020 г.*

### ИЗПОЛЗВАНА ЛИТЕРАТУРА

**Fillmore 1968:** Fillmore, C. J. The case for case. – In: Bach, E., Harms, R. (Eds.). *Universals in Linguistic Theory*. New York, 1968, 1–88.

**Fillmore 1977:** Fillmore, C. J. The case for case reopened. – In: Cole, P., Sadock, J. (Eds.). *Syntax and Semantics: Grammatical Relations*. Vol. 8, New York, 1977, 59–81.

**Fillmore 1982:** Fillmore, C. J. Frame semantics. – In: *The Linguistic Society of Korea, Linguistics in the Morning Calm*. Seoul, 1982, 111–137.

**Fillmore, Baker 2010:** Fillmore, C. J., Baker, C. F. A frames approach to semantic analysis. – In: *Oxford Handbook of Linguistic Analysis*. Oxford, 2010, 313–341.

**Johnson, Fillmore 2000:** Johnson, C., Fillmore, C. J. The FrameNet tagset for frame semantic and syntactic coding of predicate-argument structure. – In: *Proceedings of the Applied Natural Language Processing Conference (ANLP 2000)*, 2000, 56–62.

**Koeva 2020:** Koeva, S. Semantic relations and conceptual frames. – In: Koeva, Sv. (Ed.) *Towards a Semantic Network Enriched with a Variety of Semantic Relations*. Sofia, 2020, 7–20.

**Koeva 2021:** Koeva, S. Towards Expanding WordNet with Conceptual Frames. – In: *Proceedings of the 11th Global WordNet Conference, 2021*, 182–191.

**Koeva, Doychev 2022:** Koeva, S., Doychev, E. Ontology Supported Frame Classification. – In: *Proceedings of the Fifth International Conference Computational Linguistics in Bulgaria (CLIB 2022)*, 2022, 203–213.

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

Светла Коева  
професор, доктор  
Секция по компютърна лингвистика  
Институт за български език „Проф. Любомир Андрейчин“  
Българска академия на науките  
София, България

Svetla Koeva  
Professor, PhD  
Department of Computational Linguistics  
Institute for Bulgarian Language Prof. Lyubomir Andreychin  
Bulgarian Academy of Sciences  
Sofia, Bulgaria