

# КОНЦЕПТУАЛНИ ФРЕЙМОВЕ ПРИ НЯКОИ ГЛАГОЛИ ЗА КОМУНИКАЦИЯ

*ХРИСТИНА КУКОВА (СОФИЯ, БЪЛГАРИЯ)*

**Abstract.** The present study is focused on conceptual frames and includes consistent regularities within the process of their creation. One of the important tasks of the analysis is to outline a verb's selective restrictions by specifying the acceptable combinations of predicates represented as synset literals, on the one hand, and noun synsets that can occupy the position of core frame elements, on the other. Therefore, it is necessary to make a more detailed analysis of the participants in the situation as well as of their semantics. The main principles on which the work is based are indicated. The observations are focused on communication predicates and the corresponding synsets they are represented with in Bulnet, along with the frames they are attributed from FrameNet. The study aims to illustrate how each frame element that the predicate attaches can be assigned a class, a subtree, or a synset from WordNet. In this way, the scope of the core frame elements (the main participants in the described situation) is limited by specifying the semantic restrictions and the syntactic realization of each of the elements. As a result of the study, more detailed semantic restrictions have been defined in relation to the nouns with which the considered communication subclasses can be combined.

**Keywords:** conceptual frames, communication verbs, selectional restrictions

## I. ВЪВЕДЕНИЕ

Изследването, представено в настоящата статия, е съсредоточено върху предикати за комуникация и обхваща елементи от Булнет и Фреймнет. Информацията за глаголите се надгражда с допълнителни семантико-граматични наблюдения, специфични за всяко от значенията, в които се срещат, а обобщенията са отразени в Булфрейм. Описанието на глаголните единици включва данни от синонимните им множества в Булнет, информация за морфологичните и синтактичните им характеристики, както и фреймова част, основаваща се върху концептуалната структура на анализирания предикати. За целта синонимните множества, към които принадлежат, се съпоставят с фреймовете от Фреймнет. Цели се да бъдат прецизирани концептуалните елементи, които получават синтактична реализация като аргументи или адюнкти на глагола, както и селективните ограничения, налагани върху тези елементи. Част втора представя трите ресурса – Булнет, Фреймнет и Булфрейм, които се използват при това изследване, техните основни характеристики и функционалности. Част трета представлява методология на изследването, теоретична обосновка и връзка с досегашни изследвания по темата, които продължава и надгражда. Част четвърта очертава общата типология на изследваните глаголи, дефинира подгрупи и описва основните фактори и методи за специфициране на аргументите. Пета част обобщава резултатите от изследването.

В нея са представени общите ограничения на фреймовите елементи, които разглежданите групи предикати присъединяват. В заключителната част се обобщават изводите от досегашната работа и се маркират насоки за бъдещи изследвания.

## II. ЛИНГВИСТИЧНИ РЕСУРСИ

Уърднет (Miller 1995; Fellbaum 1998) е семантична база данни, която представя лексикалносемантичното познание под формата на мрежовидна структура. Възлите на тази структура представляват синонимни множества, а мисловните линии, с които са свързани, описват различни семантични, морфо-семантични, деривационни, екстратематични и други релации. Хиперо-хипонимната релация е основата, с която се характеризира структурата на Уърднет. Тя онагледява връзката между общото и частното и означава наследяване на семантични характеристики в посока надолу по дървото. В рамките на изследването селективните ограничения са представени именно като дървовидни структури с общ хипероним (клони или поддървета). Булнет (Коева 2021a) е една от идентичните на Принстънския уърднет семантични мрежи, изградена за български език. Съответните едно на друго синонимни множества в различните езици са свързани посредством уникален идентификационен код, което прави наблюденията върху примерите подходящи за обработка и от човек, и от машина; в голяма степен езиково независими и приложими за различни бази.

Фреймнет (Baker et al. 1998; Baker 2008) представлява концептуално знание, организирано в йерархична структура. Основните единици на тази структура се наричат фреймове и описват определен вид обекти, ситуации или събития заедно с техните компоненти, известни като фреймови елементи. Различават се три вида фреймови елементи – ядрени, периферни или екстратематични (Ruppenhofer et al. 2016), а тяхната комбинаторика прави фреймовете уникални. Ядрените фреймови елементи са във фокуса на внимание за целите на настоящото изследване. Те онагледяват концептуално необходимите за дадена ситуация компоненти и затова са от съществено значение при описанието на предикатите и обкръжението им. Периферните понятия като време, място, начин и др. остават извън обхвата на проучването, защото могат да бъдат заемани от представители на разнородни класове.

Булфрейм (Коева, Douchev 2022) е специално разработена система за създаване, корекция и валидиране на концептуални фреймове. „Концептуалният фрейм дефинира уникален набор от синтагматични релации между глаголни синсети, определящи фрейма, и именни синсети, изразяващи фреймовите елементи“ (Коева, Douchev 2022: 203). По този начин той представлява комбинация от семантичните данни на Булнет и Фреймнет, като ги надгражда и прецизира. Онтологичното представяне на класовете от съществителни имена в Булнет дава възможност да се дефинира точният набор от съществителни

телни, които са подходящо обкръжение за дадено значение на българските глаголни литерали. Този основен подход представлява съществена разлика при описанието на глаголите в Булфрейм и Фреймнет.

### III. ТЕОРЕТИЧНА ОБОСНОВКА

За теоретичен контекст на изследването служи семантиката на фреймовете на Филмор, според която лексикалното значение предпоставя структурата и ограниченията на синтактичната реализация. Глаголните аргументи, наричани фреймови елементи, отразяват участниците в дадена ситуация, а съответно и фрейма, описващ глагола. Така значенията на глагола се проявяват чрез фреймовете, в които участват, както и чрез модели от комбинации на аргументи, които присъединяват.

Вече повече от десетилетие се правят различни опити за комбиниране и съвместяване на информацията от Уърднет и Фреймнет, което да използва и същевременно да подсили индивидуалните им качества. Лапара и Ригау предлагат изцяло нов ресурс – УърдФреймНет (Laparra, Rigau 2010), а преди тях опити за съотнасяне правят също Фелбаум и Бейкър (Fellbaum, Baker 2008), Тонели и Пигин (Tonelli, Pighin 2009), както и екипът на Ферандез (Ferrandez et al. 2010).

Повечето от тези проучвания на лексикални единици не дават особено прецизни резултати, тъй като често не вземат под внимание релационната структура и посоката на наследяване на семантични данни в ресурсите. Съотнасянията разчитат на еднозначно свързване между лексикални единици от Фреймнет и литерали от синонимните множества в Уърднет. По-богато кодираната лексикална информация в структурата на Уърднет прави тази процедура трудна за изпълнение.

С. Лесева и др. (Лесева и др. 2018) правят по-късен опит в същата посока, като използват релационните структури на двата ресурса и свързват фреймовете със синонимни множества, вземайки предвид наследяването на концептуалните признаци в низходяща посока по дърветата.

С. Коева отбелязва необходимостта от формално описание на синтагматичните отношения в Уърднет, като въвежда понятието **концептуален фрейм**, с което дефинира множество от глаголи, проявяващи еднозначни синтагматични отношения със съществителни (Коева 2021b: 182).

Формулирането на концептуални фреймове изисква аналитичното съпоставяне на структурно-семантичните данни от Фреймнет и Булнет, като си поставя за цел дефиниране на оптимално прецизни сведения за всеки глаголен литерал и подходящо поддърво или поддървета от именни синонимни множества, за които ръчно е потвърдено, че могат да заемат позицията на даден аргумент. Така една от основните задачи при описанието, след като се обобщят резултатите по отношение на съответния семантичен фрейм, е

представянето на семантичните характеристики на фреймовите елементи, или т. нар. ограничения, и ядреният им статус.

Този процес протича в няколко стъпки:

1) Проверка на глаголния литерал и наличието му в рамките на един или няколко синонимни множества в Булнет;

2) Избор на подходящия фрейм, а при липсата на такъв – създаване на нов фрейм в съответствие с лексикалното значение на глагола;

3) Описание на лексикалната, морфологичната, синтактичната и фреймовата информация за всеки отделен глаголен литерал;

4) Избор на граматично и семантично подходящи кандидати за аргументи на предиката под формата на синонимни множества, които могат да запълнят позицията на всеки един ядрен фреймов елемент.

Направен е опит семантичните ограничения да бъдат представени с помощта на синонимни множества от Булнет, като се посочва възможно най-високият възел от дървото, който отговаря на изискванията на глагола относно потенциалната реализация на аргументите му.

Дефинирането на семантичните ограничения за всеки ядрен фреймов елемент специфицира различни концептуални модели, които обосновават подялбата на подгрупи или подфреймове.

С цел максимално прецизно представяне на концептуалния фрейм структурата му се състои от лексикална, морфологична и фреймова (семантична и синтактична) информация.

**Лексикалната информация** включва данни, които в повечето случаи са автоматично извлечени от Булнет/Уърднет: **основна форма (лема)** – основните форми са приписани автоматично и в процеса на работа се проверяват и при необходимост – коригират, **структура на лемата** (показва дали лемата е единична дума или вид съставна глаголна лексикална единица), **част от речта** (със стойност **v** – глагол), наследена автоматично от Уърднет, **междуетиков индекс в Уърднет**, с помощта на който даден глагол наследява информацията, свързана със съответното синонимно множество в семантичната мрежа, **индекс на значението**, който заедно с лемата еднозначно определя графичната единица като изразител на дадено лексикално значение; например *разказвам:1* и *предупреждавам:2*), **тълковна дефиниция** (описваща значението по начин, който го отличава еднозначно от останалите значения на същата дума), **семантичен клас на синонимното множество** (анализирани са само глаголи от класа *verb.communication*), **бележки** (информация за стилистичната маркираност: преносно значение; принадлежност към книжовната, разговорната, диалектната, терминологичната или поетичната лексика, сленга; маркираност по други показатели: нежелателна употреба, рядка употреба и др.); релации към други синонимни множества.

**Морфологичната информация** се въвежда в Булфрейм и включва сведения за **персоналността** (личен, безличен, третоличен, множественоличен),

**вида на глагола** (свършен, несвършен, двувидов, само свършен вид, само несвършен вид) – наследява се от семантичната мрежа Уърднет и **транзитивността** (преходен, непреходен) – стойността на категорията се определя въз основа на конкретното значение и се отразява в дефиницията на синонимното множество.

**Фреймовата информация съдържа: име на семантичния фрейм** (автоматично се приписва името на семантичния фрейм, ако има подходящ, дефиниран във Фреймнет, ако няма подходящ, се създава нов, специфичен за българския език; тъй като в рамките на един семантичен фрейм се определят множество различни концептуални фреймове, всеки концептуален фрейм се дефинира от глагола, към който принадлежи), **начин на създаване** (от Фреймнет, използван без промени; от Фреймнет с модификации; семантичен фрейм, който няма съответствие в оригиналния Фреймнет), **дефиниция** (автоматично се приписва дефиницията на семантичния фрейм от Фреймнет, ако е идентифициран подходящ), **релации към други фреймове** (автоматично се приписват някои от семантичните релации между фреймове, които имат най-голямо значение за структурирането на глаголната лексика), **ядрени елементи на фрейма** – название, синтактична категория, синтактична функция, селективни ограничения.

#### IV. Глаголи за комуникация

В тази част представяме извадка от предикати, означаващи взаимодействие и размяна на съобщения от сферата на писменото или устното общуване. В Булнет те са обединени като представители на семантичния клас *verb. communication*, а във Фреймнет включват лексикални единици от фрейма **Комуникация** (Communication), както и от фреймове, които влизат с него или едни с други в релациите наследяване (Inherits from) или използване (Uses). За улеснение използваме термина „глаголи за комуникация“, когато имаме предвид глаголи, поне едно от значенията на които принадлежи на този клас. Анализирани са глаголи от фреймовете за **Заявление** (Statement), **Казване** (Telling), **Изразяване на пряка и непряка преценка** (Judgment\_direct\_address, Judgment\_communication).

#### V. РЕЗУЛТАТИ

1. Концептуалната структура на глаголите с приписан фрейм Заявление и Казване представя волевия акт на предаване на Съобщение по определена Тема от Говорещия към Адресата. Тези глаголи често селектират Съобщение или Тема, като двата фреймови елемента рядко се срещат заедно в рамките на едно изречение. Те обичайно споделят общи селективни ограничения, а именно когнитивно Съдържание, което може да бъде представено и от цяла пропозиция. Говорещият е волево одушевено същество или квазиодушевена

същност, а Адресатът е Познавателен субект. Средствата за предаване на Съобщението, т.е. различно оборудване или начини за предаване на информацията, също са част от ядреното концептуално описание на тази група предикати.

### 1.1. Заявление (Statement)

В рамките на този фрейм Говорещият се обръща със Съобщение към Адресата, като използва езикови средства или различни по вид сигнали. Вместо Говорещия (или заедно с него) може да бъде запълнена позицията Средство. Темата може да бъде изразена на мястото на Съобщението. Представители на този фрейм са глаголи от синонимните множества {**обсъждам:1; обсъдя:1; разисквам:1; дискутирам:1;**}, {**заявявам:3; заявя:3; твърдя:4**} и др. По-долу са представени ядрените фреймови елементи и потенциални кандидати за запълване на съответните аргументни позиции<sup>1</sup>.

- Говорещ: Волеви одушевени субекти (човек: 4 U социална група:1)
- Съобщение: Съдържание (информация:2)
- Тема: Съдържание (реална същност:1)
- Средство: Среда или информационен канал (комуникационна система:1

U послание:1)

(1) *[Българският консул във Валенсия Пламен Георгиев]ГОВОРЕЩ заяви [по телефона]СРЕДА [за БНТ]АДРЕСАТ, [че ще коментира новината...]СЪОБЩЕНИЕ.*

### 1.2. Казване (Telling)

Фреймът описва ситуация, при която Говорещият се обръща към Адресата със Съобщение, което може да е индиректно назована Тема. Адресатът при тази група глаголи е ядрен фреймов елемент, което отличава този фрейм от други сходни на него. Вместо Говорещия (или заедно с него) може да се запълни и позицията Средство. Синонимните множества {**споделям:2, споделя:2**}, {**казвам:7, кажа:7, съобщавам:1, съобщя:1**} съдържат характерни за този тип глаголи. Ядрените фреймови елементи и възможните синонимни множества за съответните позиции са:

- Говорещ: Волеви одушевени субекти (човек: 4 U социална група:1)
- Съобщение: Съдържание (информация:2)
- Тема: Съдържание (реална същност:1)
- Средство: Среда или информационен канал (комуникационна система:1

U послание:1)

- Адресат: Познавателен субект (човек: 4 U социална група:1 U животно:1)

(2) *[И той]ГОВОРЕЩ [ни]АДРЕСАТ сподели [по телефона]СРЕДСТВО, [че от минаващи злодеи пострадал негов съсед]СЪОБЩЕНИЕ.*

2. Фреймовете **Изразяване на (непряка) преценка** и **Изразяване на пряка преценка** се отнасят до глаголи, изразяващи оценка на дадена същност

<sup>1</sup> Поради ограничения обем в рамките на това изследване е дискутиран семантичният обхват на фреймовите елементи. Структурните особености ще бъдат във фокуса на следващ материал по темата.

вместо информация. И в двата фрейма Комуникаторът прави изявление относно даден Оценяван обект, който може да бъде Познавателен субект, но също така и всяко когнитивно Съдържание. Преценката се отправя към Адресата и в двата фрейма, като той съвпада с Оценявания обект във фрейма за Изразяване на пряка преценка. Преценката може да засяга определена Тема или да е споделена с помощта на Средство. В случаите, в които Комуникаторът не е експлицитно изразен, може да се използва Изразител, т.е. част от тялото или действие на част от тялото (поглед, кимване, жест и т.н.), които да предават желаната оценка. Фреймовият елемент Причина е съществена част от концептуалното описание на глаголите за преценка.

### **2.1. Изразяване на (непряка) преценка (Judgment\_communication)**

В този фрейм Комуникаторът съобщава преценката си за Оценявания обект на даден Адресат. Преценката се реализира в положителен план (напр. хваля) или в отрицателен (напр. критикувам). Представители на този фрейм са глаголи от синонимни множества като: **{осмивам:2, осмеея:2}**, **{възпявам:1, възпея:1}**. Ядрените фреймови елементи са изразени, както следва:

- Комуникатор: Волеви одушевени субекти (човек: 4 U социална група:1)
- Тема: Съдържание (реална същност:1)
- Адресат (неядрен ФЕ): Познавателен субект (човек: 4 U социална група:1 U животно:1)
- Изразител: Част на тялото U Човешко действие (част на тяло:1 U изражение:1 U действие:2)
- Оценяван обект: Познавателен субект U Конкретно U Съдържание (реална същност:1)
- Причина: Съдържание (реална същност:1)
- Средство: Среда или информационен канал (комуникационна система:1 U послание:1)

(3) *...откакто [класната]КОМУНИКАТОР [я]ОЦЕНЯВАН ОБЕКТ похвали [пред целия клас]АДРЕСАТ [за бележките ѝ]ПРИЧИНА, всеки искаше да седи на един чин с нея и заедно да си учат уроците.*

### **2.2. Изразяване на пряка преценка (Judgment\_direct\_address)**

В рамките на този фрейм Комуникаторът отправя преценката си относно Оценявания обект към Адресата, като в този случай Оценяваният обект и Адресатът съвпадат. Преценката се прави по определена причина или Тема. Фреймът включва глаголи от следните синонимни множества: **{укорявам:3; укоря:3; предупреждавам:3; предупрежда:3; упреквам:2; упрекна:2;}**, **{благодаря:2}** и ядрените фреймови елементи:

- Комуникатор: Волеви одушевени субекти (човек: 4 U социална група:1)
- Тема: Съдържание (реална същност:1)
- Адресат: Познавателен субект (човек: 4 U социална група:1 U животно:1)
- Изразител: Част на тялото U Човешко действие (част на тяло:1 U изражение:1 U действие:2)

- Оценяван обект = Адресат: Познавателен субект (човек: 4 U социална група:1 U животно:1)
- Причина: Съдържание (реална същност:1)
- Средство: Среда или информационен канал (комуникационна система: 1 U послание:1)

(4) *Най-вече понеже [отговорният другар]КОМУНИКАТОР [ме]ОЦЕНЯВАН ОБЕКТ = АДРЕСАТ поздравя [за изказването]ПРИЧИНА.*

## VI. ИЗВОДИ И БЪДЕЩА РАБОТА

Изследването очертава както общата стратегия и очакваните закономерности при извеждане на семантичните ограничения, така и някои специфики, свързани с отделни подгрупи глаголи. Предстои работа върху по-детайлно презициране на хипонимите, които могат да заемат позицията на фреймовия елемент при всяка глаголна лексикална единица. Имплементирането им в Булфрейм несъмнено ще даде нови насоки за изследователи и потребители.

## БЛАГОДАРНОСТИ

*Изследването е извършено по проект „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.*

## ИЗПОЛЗВАНА ЛИТЕРАТУРА

**Лесева и др. 2018:** Лесева С., Стоянова, И., Кукова, Х., Тодорова, М. Интегриране на субкатегоризационна информация в релационната структура на УърдНет. – Български език, 65, 2018, № 2, 13–41. [Leseva, S., Stoyanova, I., Kukova, H., Todorova, M. Integrirane na subkategorizatsionna informatsiya v relatsionnata struktura na WordNet. – Balgarski ezik, 65, 2018, № 2, 13–41.]

**Baker 2008:** Baker, C. F. FrameNet, present and future. – In: Webster, J., Ide, N., Chengyu Fang, A. (Eds.). The First International Conference on Global Interoperability for Language Resources. Hong Kong, 2008, 12–17.

**Baker et al. 1998:** Baker, C. F., Fillmore, C., Lowe, J. The Berkeley FrameNet Project. – In: Proceedings of COLINGACL, 1998, Canada, 86–90.

**Fellbaum 1998:** Fellbaum, C. (Ed.). WordNet: An Electronic Lexical Database. Cambridge, MA, 1998.

**Fellbaum, Baker 2008:** Fellbaum, C., Baker, C. F. Can WordNet and FrameNet be Made “Interoperable”? – In: Webster, J., Ide, N., Chengyu Fang, A. (Eds.). The First International Conference on Global Interoperability for Language Resources. Hong Kong, 2008, 67–74.

**Ferrandez et al. 2010:** Ferrandez, O., Ellsworth, M., Munoz, R., Baker, C. F. Aligning FrameNet and WordNet Based on Semantic Neighborhoods. – In: Proceedings of the 7th Conference on Language Resources and Evaluation (LREC 2010), 17–23 May 2010, Valletta, Malta, 2010, 310–314.

**Koeva 2021a:** Koeva S. The Bulgarian WordNet: Structure and specific features. – In: Papers of BAS. Humanities and Social Sciences, vol. 8, 2021, No. 1, 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, 2021, 182–191. <<https://aclanthology.org/2021.gwc-1.21>>

**Koeva, Doychev 2022:** Koeva, S., Doychev, E. Ontology Supported Frame Classification. – In: Proceedings of the 5th International Conference on Computational Linguistics in Bulgaria (CLIB 2022), 2022, 203–213.

**Laparra, Rigau 2010:** Laparra E., Rigau, G. eXtended WordFrameNet. – In: Proceedings of the 7th Conference on Language Resources and Evaluation (LREC 2010), May 17–23, Valletta, Malta, 29 May–2 June 2010, Las Palmas, Spain, 1214–1219.

**Miller 1995:** Miller, G. A. WordNet: a lexical database for English. – In: Communications of the ACM, vol. 38, 1995, No. 11, 39–41.

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

**Tonelli, Pighin 2009:** Tonelli, S., Pighin, D. New Features for Framenet – Wordnet Mapping. – In: Proceedings of the 13th Conference on Computational Natural Language Learning (CoNLL'09), Boulder, USA, June 2009, 219–227.

Христина Кукова  
главен асистент, доктор  
Секция по компютърна лингвистика  
Институт за български език „Проф. Любомир Андрейчин“  
Българска академия на науките  
София, България

Hristina Kukova  
Chief Assistant Professor, PhD  
Department of Computational Linguistics  
Institute for Bulgarian Language  
Bulgarian Academy of Sciences  
Sofia, Bulgaria