

СЕМАНТИЧНИ РЕЛАЦИИ: ТЕОРЕТИЧНИ И ПРИЛОЖНИ АСПЕКТИ

ИВЕЛИНА СТОЯНОВА, МАРИЯ ТОДОРОВА

ИНСТИТУТ ЗА БЪЛГАРСКИ ЕЗИК „ПРОФ. Л. АНДРЕЙЧИН“ ПРИ БАН
iva@dcl.bas.bg, maria@dcl.bas.bg

SEMANTIC RELATIONS: THEORETICAL AND PRACTICAL ASPECTS

IVELINA STOYANOVA, MARIA TODOROVA

INSTITUTE FOR BULGARIAN LANGUAGE *PROF. L. ANDREYCHIN*
BULGARIAN ACADEMY OF SCIENCES
iva@dcl.bas.bg, maria@dcl.bas.bg

The paper presents an overview on the theoretical aspects of current descriptions of semantic relations, as well as the practical aspects of various systems of semantic relations used in wordnets for different languages in view of the possibilities to extend the Bulgarian lexical-semantic wordnet BulNet with new semantic relations.

Keywords: semantic relations, WordNet, BulNet

1. Въведение

Статията има за цел изследването на различни описания на семантичните отношения между думите с оглед на разширяването на Българската лексикално-семантична мрежа БулНет с нови семантични релации. Представените теоретични постановки за същността на семантичните релации и разгледаните приложения на семантичните отношения нямат претенции за пълна изчерпателност. Релациите са описани от гледна точка на приложението им в Българския уърднет като представител на лексикално-семантичните мрежи с оглед на изследване на възможностите за разширяване на цялостната му логическа структура чрез дефиниране на нови типове релации и приписването им за български език. Наред с теоретичния преглед на описанието на семантичните релации са представени и някои междуезикови паралели с уърднети на различни езици с оглед на възможностите за пренасяне на семантични релации с цел обогатяване на БулНет.

1. Семантични отношения

1.1. Същност на семантичните релации

Мисловната дейност се характеризира със стремежа на индивида към изграждане на единна представа за заобикалящия го свят. Според Пернишка „при възприемане на действителността чрез сетивата се извършва кате-

горизация на човешкия опит, при което получената от човека информация се подрежда и систематизира под формата на мисловни същности, резултат от познанието – концепти (понятия)“ (Пернишка/Pernishka 2014: 405). Каплан и Херман (Caplan and Herrmann 1993) посочват, че логическата структура на мисленето и умението на човека да разбира и разсъждава се основават на разнообразието от семантични отношения и техните свойства. Концептуалните структурни отношения са част от начина, по който организираме и изграждаме своето знание и разбиране за света. То зависи от способността ни да възприемаме и характеризираме отношенията между понятията и да създаваме концептуални класове и връзки помежду им.

В книгата си „Семантични структури“ (Jackendoff 1990) Джакендоф дефинира лексикално-концептуалните структури на езиковите изрази като абстрактни мисловни репрезентации, чрез които носителите на даден език организират ситуациите в света. Ако концептите се разглеждат като основни строителни елементи на концептуалната структура, то отношенията между тях, наричани релации, са свързката, която ги държи заедно.

Ку и На (Khoо and Na 2007) посочват, че взаимоотношенията между концептите се определят чрез едновременно идентифициране на семантична връзка и набор от участници в нея, като всеки участник има своя роля във взаимоотношенията. Концептите са неразделна част от релациите, които не могат да съществуват сами по себе си, а трябва да свързват две неща. Релационните структури често се представят като тройка: [концепт₁] → (връзка) → [концепт₂]. Всяка релация свързва две или повече позиции, които трябва да бъдат запълнени, като тяхната селекция е ограничена и позициите може да бъдат запълвани само от участници, принадлежащи към определена семантична категория. Наличието на релации на базовото ниво на представяне на знания позволява генериране на единици на по-комплексно ниво. Семантичното съдържание на релациите определя природата на единиците от по-високо ниво. Важността на релациите не е ограничена само до това как представяме знанието, а се състои и в това как го използваме, когато разсъждаваме. Релационното познание е ключ за начина, по който разсъждаваме за света. Например ако знаем, че А и Б са свързани като *клас* и негова *проява*, ние ще разсъждаваме за тях по напълно различен начин от това, което бихме мислили, ако знаехме, че те са свързани като *причина* и *следствие*.

Разбирането за богатството от релации ни кара да изследваме ролята им в човешкото възприятие на света, от една страна, и ролята им в по-формални системи, които представят и боравят със знание, от друга. Според Спелман и др. (Spellman et al. 2001) концептуалните отношения и ролята, която те налагат върху участниците в ситуацията, са от основно значение за когнитивни задачи като разбиране, разсъждение, заключение, разрешаване на проблеми, аналогични разсъждения и използване на метафори и др.

Грийн (Green 2002) изследва ролята на релационната структура за човешкото мислене, възприятия, разсъждаване и въображение от перспективата на когнитивната семантика. Тя твърди, че списъкът на семантичните връзки включва както затворен брой отношения на йерархичност и еквивалентност, така и отворен набор от отношения. Всеки нов глагол носи потенциал за въвеждане на нова концептуална връзка. Всяка връзка между обекти, процеси и др. може да бъде отразена в езиците между съответните понятия. Например *любов* е връзка между хора, „Том“ и „Клеър“ могат да са индивидуални понятия, свързани със семантичната връзка *любов*. Различните тематични области непрекъснато развиват нови видове семантични отношения.

Освен отношенията в света, семантичните релации могат да отразяват отношенията в езика, включително отношенията между обектите и техните символи. Признаковият подход в лингвистиката, въведен от Якобсон и Хал (Jakobson and Halle 1956), е свързан с идеята за минималните различителни елементи, или диференциалните признаци. Семантичните изследвания на Сапир (Sapir 1921) и Йелмслев (Ельмслев/Нjelmslev 1960) за фигурите на съдържанието дават тласък на компонентния анализ на значенията.

1.2. Същност на релациите

Релациите изпълняват ключова роля при представянето на психологични, лингвистични, компютърни и други знания. Много от системите за представяне на знания се основават на експлицитни разлики между единиците, отразяващи се в отношенията помежду им, каквито например са възлите и дъгите при семантичните мрежи. Най-често използваните семантични отношения напомнят списъка от категории, посочени като универсалии в класификацията на понятията на Аристотел (Аристотел/Aristotel 1992): *субстанция, количество, качество, отношение, място, време, положение, притежание, действие, засегнатост/подложеност (на действие)*. Те в голяма степен се припокриват с общоупотребяваните граматически категории и представляват абстракции или обобщения. Примерът с релацията *любов* може да се разглежда като частен случай на релацията *засегнатост*.

Въпреки че броят на семантичните отношения изглежда неограничен, в повечето случаи на практика се използва ограничен брой общи групи. В рамките на Уърднет се използва затворен клас от езикови взаимоотношения, които създават неговата структура, например хипонимия, меронимия, синонимия, антонимия и други, т.е. еквивалентни и минимални йерархични връзки. Взаимоотношенията, изразяващи асоциативни връзки, са отворен клас, например връзките между глагол и деривационно свързаните с него думи. Тъй като е отворено, членството в класа никога не може да бъде напълно изброено: когато се създава нов глагол, възниква потенциалът за въвеждане на нова концептуална връзка.

Определянето на свойствата на релациите включва признаци, свързани с отехния обхват: идентифициране на нивото на участниците (като класове или като конкретни единици); брой на участниците във взаимоотношени-ята и тяхната кардинална или транзитивна насоченост.

1.3. Признаци на семантичните релации

Семантичните отношения възникват на всички езикови нива – между думи, фрази, прости изречения, изречения и по-големи текстови сегменти, както и между документи и множество документи. Мърфи (Murphy 2003) изброява следните общи свойства на лексикално-семантичните отношения:

1. Продуктивност – потенциал за създаване нови релации.
2. Бинарност (двуместност) – потенциал за свързване на двойки елементи. Някои релации, като например антонимията, са двоични. Дадена дума може да има най-много един истински антоним. Други семантични отношения, като например синонимията, могат да се отнасят за набор от думи (т.е. една дума може да има много синоними).
3. Променливост – потенциал за промяна на връзките между думите в зависимост от смисъла на използваната дума и контекста.
4. Прототипност и каноничност – еднозначен каноничен статус на конкретни релации (например антонимите).
5. Полусемантичност – потенциал за идентифициране на дадена релация с помощта на несемантични признаци като принадлежност към дадена граматична категория; взаимна селективност; сходство в морфологичната форма.
6. Неизброимост – семантичните отношения са отворен клас и те не могат да бъдат изброени или изчерпани.
7. Предсказуемост – семантичните отношения следват определени общи модели и правила.
8. Универсалност – едни и същи типове семантични отношения се използват във всички езици, като на различните езици едни и същи релации свързват едни и същи концепти.

Крус (Cruse 2002; Cruse 2011), Лайънс (Lyons 1977; Lyons 1995) и Мърфи (Murphy 2003) подробно описват семантичните отношения в езика. Обобщавайки ги, Ку и На (Khuo and Na 2007) посочват следните логически свойства на семантичните релации:

- Рефлексивност: релацията R е рефлексивна, ако тя може да свърже дадена единица със себе си. Т.е. $[x] \rightarrow (R) \rightarrow [x]$ е вярно за всяко x (например релацията *част от*);
- Симетричност: релацията R е симетрична, ако двамата участници в отношението могат да заемат всяка от двете позиции. Т.е. $[x] \rightarrow (R) \rightarrow [y]$ предполага, че $[y] \rightarrow (R) \rightarrow [x]$ (например релацията *синонимия*);
- Транзитивност: релацията R е транзитивна, ако $[x] \rightarrow (R) \rightarrow [y]$ и $[y] \rightarrow (R) \rightarrow [z]$ (например *хипонимно – хиперонимната* релация);

- **Фиксираност:** релацията R е фиксирана, ако конкретен участник в отношението предопределя точно определен друг участник. Т.е. $[x] \rightarrow (R) \rightarrow [y]$ и $[z] \rightarrow (R) \rightarrow [y]$ предполага, че $x = z$ (например релацията *антонимия*).

Релациите могат да влизат във взаимоотношения помежду си. Такъв тип отношения например са *сходство* или *противопоставеност* между релации. Релацията R е обратна на релацията S, ако и двете могат да приемат една и съща двойка участници чрез обръщане на посоката на двете отношения. Т.е. $[x] \rightarrow (R) \rightarrow [y] \rightarrow (S) \rightarrow [x]$ (например релациите *родител* и *дете*). Дадена релация може да бъде подчинена или специфична проява на друга. Релациите могат да бъдат организирани в йерархия.

1.4. Типове семантични релации

Изхождайки от свойството на релациите да съпоставят концепти, Грийн (Green 2001) дефинира три основни метакласа: а) отношения на еквивалентност, представени от синонимията и антонимията; б) отношения на йерархичност, представени от хипонимията, включването и меронимията (отношение между част и цяло) и в) отношения на асоциативност, представени от връзката *причина – ефект*.

Релациите между понятия Пернишка разглежда от гледна точка на вторичната номинация, при която готово име W се пренася като название на друг обект $S1 \rightarrow D1$. Назоваването се основава на връзка от различен характер между първично и вторично означените денотати или сходство между тях. Пернишка посочва пет типа формалнологически отношения между понятията, основани на вторичната номинация: *равнозначност, изключване, подчинение (включване), пресичане и противоположност* (Пернишка/Pernishka 2014: 518). Авторката разглежда три основни типа семантично пренасяне (многозначност) – метафора, метонимия и енантиосемия, и дефинира техните подтипове:

- **Метафората** назовава качествените особености на явленията и свързва понятията с оглед на сходство, близост, аналогия по такива признаци. Нейни разновидности са непряка, косвена номинация и квалификативна номинация.

- **Метонимията** е преименуване, даване на друго съседно име. Нейни разновидности са парадигматично мотивираната метонимия и синтагматично мотивираната метонимия.

- **Енантиосемията** е проява на полисемия, при която една лексема развива две взаимно противоположни (антонимични) значения. Нейни разновидности са:

- а. **Лексикалносемантичната енантиосемия:**

- ✓ контрарна (градуална) енантиосемия за максимална/минимална степен на признак: сигурно ‘напълно възможно’ ↔ ‘малко възможно’;
- ✓ векторна енантиосемия – за противоположност по време или пространство: скоро – в миналото ↔ в бъдещето;

✓ комплементарна енантиосемия – противопоставяне по взаимнодопълващи се признаци: *Оставям нещо на масата* ('давам' = 'правя да има нещо някъде');

✓ конверсивна енантиосемия: *заемам на някого (давам) ↔ заемам от някого (вземам)*;

б. Лексикално-морфологична енантиосемия, съвместяване на две противоположни морфологични значения в една лексема (*уста, врата*);

в. Семантично-стилистична енантиосемия, при която противоположен е емоционалният признак в двете значения (положителен ↔ отрицателен): *хитър* (с положителна ↔ с отрицателна оценка).

Лангакер (Langacker 1987) групира отношенията между понятията в предикативни групи: г) номинални предикации, които означават неща и съответстват на съществителните; д) релационни предикации, които означават състояния, атемпорални отношения и процеси и съответстват на прилагателни, наречия, предлози и глаголи; е) термини и логически позиции от първи ред.

Телефсен (Thellefsen 2010) изброява следните релации:

• Активна релация – семантична връзка между две понятия А и Б, като А изразява изпълнението на операция или процес, засягащи Б;

• Антонимия – семантична връзка между понятия А и Б, като А е обратното на Б (например *студено* е обратното на *топло*);

• Асоциативна релация – семантична връзка между понятия А и Б, като А е психически свързано с Б от някого;

• Причинно-следствена релация – семантична връзка между понятия А и Б, като А е причината за Б. (например: *Скорбутът се причинява от липсата на витамин С.*);

• Омонимна релация – семантична връзка между понятия А и Б, като А и Б се изразяват с един и същ символ;

• Хипонична релация – семантична връзка между понятия А и Б, като: А е вид Б / А е подчинен на Б / А е по-тясно от Б / Б е по-широко от А. Тя е представена от релацията *хипоним – хипероним* и изразява родово отношение, родово-видова връзка или връзка на йерархично подчинение. Когато всички отношения в дадена система означават клас, към който принадлежи даден обект (например *синигерът е птица*), системата е таксономия. С релацията *произхожда от* се посочват всички конкретни видове (разновидности, хипоними) на понятието. Релацията *специфичен за* позволява да се посочи общият род (хипероним) на всички конкретни типове;

• Релация *пример за/проява на* – семантична връзка между понятия А и Б, като А е пример за Б. Обозначава семантичните отношения между общото понятие и отделните му случаи (например: *Копенхаген е пример за общото понятие столица*);

• Локативна релация – семантична връзка между понятия А и Б, като А се намира в Б. Понятието показва местоположението на нещо, обозначено от друго понятие (например: *малцинствата в България*);

- Релация *част – цяло* (меронимна) – семантична връзка между две понятия А и Б, като А е по-тясно от Б, а Б е по-широко от А. Означава взаимоотношение между цялото и неговите части. Нейни прояви са: меронимията (А е част от Б) и холонимията (Б съдържа в себе си А);

- Пасивна релация – семантична връзка между две понятия А и Б, като А е засегната от или е подложена на операция или процес, изразени от Б;

- Парадигматична релация (Wellisch 2000) – семантична връзка между две понятия А и Б, като А е свързано с Б по природа, явно или с оглед на конвенция (например *майка – дете; мазнини – затлъстяване; държава – столица*);

- Полисемна релация – семантична връзка между две или повече понятия А₁, А₂ и А₃, изразени с общ знак Б, като А₁, А₂ и А₃ споделят един и същ израз;

- Релация за притежание – семантична връзка между две понятия А и Б, като А е притежател, а Б е притежаваното от А;

- Релация за свързаност на термини – относителна семантична връзка между две или повече понятия. Свързаните термини могат да изразяват широк кръг релации – антагонистични отношения, активни/пасивни отношения, причинно-следствени връзки, жилищни отношения, парадигматични отношения;

- Синонимна релация – семантична връзка между две понятия А и Б, като А е еквивалентно/означава същото като Б;

- Времева връзка – семантична връзка между две понятия А и Б, като А посочва време или период от Б (например *Втората световна война, 1939–1945*);

- Тропонимна релация, дефинирана в Уърднет – семантична връзка между две понятия А и Б, като А е начин да се направи Б.

Редица теории основават универсалните семантични параметри в езика на връзката между определена ситуация и участниците в нея. Според генеративната лексикална теория (Pustejovsky 1995, 1998) значението на думите, наричано QUALIA структура, е комплекс от четири основни свойства, базирани на начина, по който хората разбират обектите и отношенията в света:

- Формално свойство (*is_a; type_of; has_location*), отличаващо обектите в рамките на по-широка категория;

- Конститутивно свойство (*part_of*), изразяващо връзката между обект и неговите съставни елементи;

- Целево свойство (*effected_by; result_of; has_function*), изразяващо целта и функцията на обекта;

- Агентивно свойство (*is_agent; is_patient*), изразяващо произхода или възникването на обекта.

Изследванията извеждат семантичните отношения от способността на думите да изразяват типови смисли при своето комбиниране, проявени

чрез лексикалната съчетаемост на думите. Такива са например описанията на предикатно-аргументните отношения, основани на синтактичните свойства и поведение (Levin 1993; Goldberg 1994 и др.), на тематичната структура (Longacre 1976; Van Valin and LaPolla 1997), на лексикалната концептуална структура (Jackendoff 1990), на рамковата семантика (Fillmore 1982). За български различните варианти на класификация на семантичните роли са представяни и прилагани от Коева (Коева/Koeva 2006, Коева/Koeva 2008, Коева/Koeva 2010), Влахова (Влахова/Vlahova 2005, Влахова/Vlahova 2006), Джонова (Джонова/Dzhonova 2005, Джонова/Dzhonova 2008), Ницолова (Ницолова/ Nitsolsova 2005) и др.

Според теорията за лексикалните функции (Mel'čuk and Žolkovskij 1970; Mel'čuk 1996) съчетателните възможности на дадена лексикална единица L и потенциалът ѝ за изразяване на определени значения се представя чрез нейните лексикални корелати и чрез механизма на математическите функции. Лексикалните функции са универсални и могат да описват едновременно семантични деривации и ограничения на взаимната селекция във всеки естествен език и това им позволява да представят езиково специфични значения, включително и флективни и словообразователни значения. Лексикалните корелати са множеството $\{L_i'\}$, състоящо се от повече или по-малко близки по значение лексикални единици, обединени от семантичната си връзка с дадена лексикална единица L_1 . Общият семантичен признак между L_1 и $\{L_i'\}$ се изразява чрез лексикалните функции. Дадена лексикална функция **LF**, свързваща лексикалните единици L_1 и множеството от стойности L_2, L_3, L_4, \dots , се представя чрез понятието за математическа функция $\mathbf{LF}(L_1) = L_2, L_3, L_4, \dots$. Функцията $f(L) = \{L_i'\}$ приписва на своя аргумент (ключова дума) – лексикалната единица L , стойността на множеството от лексикални корелати $\{L_i'\}$.

Според представянето на лексикалните релации в теорията „Смисъл \Leftrightarrow текст“ отношението на семантична деривация свързва две лексикални единици L_1 и L_2 при следните условия:

1. L_1 и L_2 изразяват приблизително еднакво значение. L_2 е близко до или произтича от значението на L_1 в даден контекст и принадлежи към същата част на речта. Ако L_2 принадлежи към част на речта, различна от тази на L_1 , то L_2 е глаголен, номинален, адективен или адвербиален корелат на L_1 .

2. Съществува поне една употреба, при която L_1 изразява значение, различно от значението на L_2 .

3. L_2 изразява елемент от ситуацията, представяна от L_1 . Такава е например връзката между глагола *убивам* и съществителните *убиец* и *жертва*, които са участниците в представяната от глагола *убивам* ситуация.

Множеството от лексикални функции е разпределено по части на речта. Функциите за интензификация *се представят чрез* адективната и адвербиална лексикална функция **Magn** (много, до голяма степен, интензивно) и представят формално колокацията *ОСТЪР контраст* съответно като

Magn (*контраст*) = *остър*. Глаголните лексикални функции представят различни групи глаголи и техните субкатегоризационни особености. Например функциите, представящи групата на лексикално-функционалните глаголи, **Operi**, **Funci Labor_{ij}**, чрез индекса и ключовата дума посочват семантико-синтактичните обекти в дълбинната структура. Лексикалната функция **Func_i**, при която външният семантичен обект (субектът в дълбинната структура) е ключовата дума L, а задължителният вътрешен семантичен обект е i-тият синтактичен обект, се представя с формулата **Func_i** (*инструкция*) = *се отнася [за]*.

Семантичните деривации са кодирани чрез парадигматични лексикални функции, а колокациите чрез синтагматични лексикални функции. Границата между тези два типа лексикални функции и между съответстващите им феномени не е изразена. Прилагат се така наречените застъпващи се стойности (*fused values*), чрез които на синтагматичните лексикални функции могат да съответстват семантичните деривации.

Теорията за лексикалните функции предлага модел, който едновременно описва семантични деривации и ограничения на взаимната селекция във всеки естествен език и това им позволява да представят езиково специфични значения, включително и флективни и словообразователни значения.

2. Семантични отношения и Българският уърднет (БулНет)

Семантичните релации намират приложение в различни сфери на компютърната лингвистика в системи за извличане на информация, машинен превод, разрешаване на многозначност и др. Едно от най-обширните практически описания на лексикалната система на езика, комбинираща в логическа мрежа концепти и свързващите ги семантични релации, е Уърднет (Fellbaum 1998; Miller 1995). Тук представяме релациите от гледна точка на приложението им в Българския уърднет като представител на лексикално-семантичните мрежи с оглед на изследване на възможностите за разширяване на цялостната му логическа структура чрез въвеждане на нови релации, използвайки теоретичната рамка на представените по-горе множества и подходи. Ще бъдат разгледани релации в чуждоезични уърднети и други ресурси с оглед на възможностите за пренасянето им в Българския уърднет и на дефинирането на нови типове релации и приписването им за български език.

Българският уърднет (БулНет, <http://dcl.bas.bg/bulnet/>) е създаден въз основа на Принстънския уърднет (PWN, <http://wordnet.princeton.edu/>). Като цяло той използва структурата на PWN, където основната единица е синонимното множество (синсет), съставено от отделни лексикални единици (литерали). С помощта на междуезиковия индекс между синсетите (Interlingual Index, ILI) се запазва съотнасянето както с PWN, така и с други уърднети, базирани на PWN, което разширява приложенията на БулНет в междуезикови изследвания. В PWN са дефинирани и набор от релации

между синсетите и/или литералите на различни езикови нива – семантично, лексикално, словообразователно, граматическо, които са пренесени и в БулНет. Освен това в БулНет са отразени и езиково специфични понятия и релации, които отразяват особеностите на българската лексикално-семантична система.

Коева (Коева/Коева 2007; Коева/Коева 2014) групира релациите в Уърднет в два типа: лексикални – между литерали, и концептуални – между синонимни множества. От своя страна „лексикалните релации могат да бъдат семантични или деривационни, а концептуалните – семантични, морфологично-семантични или екстралингвистични. Концептуалните релации се отнасят до всички литерали от двете синонимни множества, които свързват. Лексикалните релации свързват отделни думи в рамките на едно или на две синонимни множества“. Те също образуват мрежа, в която възлите са думи, а не синонимни множества, а дъгите – лексикални релации. Релациите на семантично ниво (между синсети) от PWN са пренесени автоматично в БулНет по следния начин:

(а) множеството от релации от PWN е запазено с малки изключения: някои случаи на релацията *включване* (entailment) са представени като *има/е подсъбитие* (has subevent/is subevent) (най-вече в случаите на темпорално включване на едно събитие в друго); синсетите с деривационни релации между литерали са свързани с *eng_derivative*, ако деривацията съществува само в английски, или с *bg_derivative*, ако има и деривация на български;

(б) Релациите от PWN са валидирани от експерт; ако дадена релация е нерелевантна за български, това е означено в бележка към българския синсет. Приема се, че две понятия могат да бъдат свързани само с една семантична релация от даден тип, затова бяха направени няколко проверки за непротиворечивост върху 450 синсета, на които бяха приписани повече от една морфосемантична релация. В резултат 290 релации бяха променени, 100 отстранени и 255 валидирани (Коева et al. 2016);

(в) Добавени са нови деривационни релации между синсети глаголи и съществителни, които също са ръчно верифицирани (Piasecki and Koeva 2017).

3. Възможности за използването на многоезични ресурси за разширяване на БулНет с нови релации

В контекста на направеното описание на съществуващите релации и тяхната реализация в различни системни отношения тук искаме да представим практически ориентиран сравнителен анализ между семантичните релации, вече установени в Българския уърднет, и релациите в някои други чуждоезични уърднети. Основната цел на анализа е да бъдат представени възможностите за заимстване на информация чрез междуезикови паралели – както чрез пряко пренасяне на релации между езиците, така и чрез използване на информацията за създаване на нови релации или допълване на съществуващата вече информация.

Основните принципи при междуезиковите анализи и пренасянето на информация между лексикално-семантични мрежи на различни езици са:

1. Осигуряване на системност и консистентност;
2. Верификация на съществуващи единици – както синсети, така и релации;
3. Автоматично разширяване на Уърднет с нови синсети и релации, както и дефиниране на нови типове релации;
4. Приложимост на използваната информация.

Едно от основните преимущества на лексикално-семантичната мрежа Уърднет е фактът, че около нея са създадени многобройни други уърднети за различни езици по модела на разширяване (Vossen 2005), като отделните уърднети са свързани помежду си с помощта на общ междуезиков индекс (interlanguage index, ILI), съотнасящ значенията, които са езиково независими (отделните синсети), а не лексикалните единици (отделните литерали).

Проблемите, свързани със съотнасянето на отделните уърднети, най-често се дължат на следните характеристики:

(а) Различните уърднети могат да са базирани на различни версии на PWN, а съществуващите съотнасяния между версиите невинаги са пълни и/или прецизни;

(б) Различните уърднети се поддържат в различни формати, най-често в зависимост от програмите, които се ползват за тяхното създаване и редактиране (XML с различни схеми, обикновен текстов формат, CSV формат и т.н.);

(в) Много уърднети не се разпространяват свободно и/или не позволяват промяна във формата, за да бъдат съотнесени към единен формат (Bond and Foster 2013);

(г) Макар да има едно общо ядро от универсални за всички уърднети семантични релации (*синонимия*, *хипонимия* – *хиперонимия* и др.), повечето уърднети съдържат и специфични семантични релации, различаващи се по названия, покритие и набор от стойности;

(д) дори да имат обща изходна база и общо ядро от понятия, различните уърднети добавят в различна степен нови и езиково зависими явления – както синсети (например по-конкретни понятия), така и релации (например деривационни), които не са съотнесени с PWN, а оттам не са съотнесени и с чуждоезичните уърднети. В резултат липсват съществуващи връзки между явления в близки езици (например български и другите славянски езици), когато това явление не съществува в английски.

На базата на PWN се създават редица многоезични лексикално-семантични мрежи, които целят въвеждането на единен формат на представяне на данните, с цел да бъдат приложими за междуезикови и многоезични изследвания и разработване на компютърни приложения. Такива са мрежи като Euro WordNet (Vossen 1998), BalkaNet (Tufiş et al. 2004), Multilingual

Central Repository (Gonzalez-Agirre et al. 2012), Open Multilingual Wordnet (Bond and Foster 2013) и други.

Възможностите за извличане и заимстване на информация от чуждо-езичните уърднети и други ресурси включват:

(а) пренасяне на релации от чуждоезичен уърднет в БулНет (релации, които вече са дефинирани в БулНет), включително в случаите, когато се налага съотнасяне на релациите от две отделни класификационни схеми;

(б) автоматично добавяне на семантична и/или прагматична информация от други ресурси в Българския уърднет, като например релации INSTANCE_OF;

(в) проверки за консистентност и непротиворечивост, както и верификация на вече установени релации в Уърднет;

(г) дефиниране на нови типове релации и приписването им за български език.

4. Преглед на типовете релации с оглед на междуезиковото им пренасяне

Автоматичното пренасяне на релации е продължение на така наречения метод на разширяването за създаване на уърднети (Piasecki and Koeva 2017), при който се използва концептуалната структура на PWN и се съблюдава принципът за запазване на структурата. Това позволява сигурното пренасяне на семантичните отношения на концептуално равнище между единиците от различни езици и по този начин се постига обогатяване със семантични релации. Релациите биват:

- Релации между синонимни редове от съществителни: синонимия, антонимия, хипонимия/хиперонимия и меронимия/холонимия;
- Релации между синонимни редове на глаголи: тропонимия, импликация и каузалност;
- Релации между различни части на речта: морфосемантични връзки; например наблюдавам (глагол), наблюдателен (прилагателно), наблюдение, наблюдаване (съществително);
- За двойките съществително – глагол може да бъде специфицирана семантичната роля на съществителното с оглед на глагола: {спалня} е локация спрямо {спя}; {гравьор} е агент на {гравирам}, докато {гравюра} е резултат.

4.1. Деривационни релации и семантичното им съдържание

В последните години се обръща значително внимание на нови начини за извличане на семантична информация за езиците с богата словообразователна система (морфологично богати езици), каквито са славянските и балканските езици (Bilgin et al. 2004; Koeva 2008; Mititelu 2012), и пренасянето на семантичните релации от един език в друг.

Морфосемантичните релации (Fellbaum et al. 2009) дават два вида езикова информация: (а) възможна езиково зависима деривационна връзка

между два литерала от различни синсети; (б) езиково независима семантична релация от определен тип между две понятия (Dimitrova et al. 2014; Leseva et al. 2015). Понастоящем в PWN (а оттам и в БулНет) са представени само морфосемантични релации между глаголи и съществителни имена. Също така не всички отношения на деривация са отразени – например липсват релации между сложни или съставни думи и простите думи, с които са свързани (например *добивам* и *въгледобив*).

Както отбелязват Лесева и др. (Leseva et al. 2015), деривационните модели, свързани с морфосемантичните релации, са многозначни. Например от 45 модела (двойки наставки на глагол и съществително), характерни за морфосемантичната релация *агент*, само 13 са еднозначни (т.е. се срещат само с релацията *агент*). За някои релации семантичният примитив на съществителното подсказва релацията. Например, наставката *-тел* с примитив *noun.person* е сигурен показател за релацията агент (например *преподавам* – *преподавател*), докато *-тел* с примитив *noun.artifact* е инструмент (например *нагривам* – *нагревател*).

Румънският уърднет отразява деривационни релации както между различни, така и в рамките на една и съща част на речта (Mititelu 2012). Деривационните модели включват както наставки, така и представки и конверсия. Откритите двойки синсети с отношения на хиперонимия и хипонимия, меронимия и холонимия се отбелязват като деривационно свързани, без да се добавят нови семантични релации. Отбелязват се морфосемантичните релации, но наборът от релации само частично се припокрива с тези на PWN, като се въвежда и неопределената релация RELATED за случаи, за които няма открита специфична релация.

В Чешкия уърднет релациите показват частите на речта, за които се отнасят, а не семантичния тип на релацията: *deriv-na*, *deriv-ger*, *deriv-dvrb* и т.н. (Pala and Hlavackova 2007). В Турския уърднет морфосемантичните релации са по-обобщени: *become*, *acquire*, *be-in-state*, *someone-with*, *something-with*, *someone-from*, *someone-without*, *something-without*, *pertains-to*, *with*, *reciprocal*, *causes*, *is-caused-by*, *cat-of*, *manner* (Bilgin et al. 2004).

Многоезичният централен архив (Multilingual Central Repository, MCR) е проект за взаимосвързани уърднети на шест езика: английски, испански, каталонски, баски, галисийски и португалски. MCR е обогатен и със семантично анотирани дефиниции, както и с тематични области, приписани на синсетите (Gonzalez-Agirte et al. 2012). Уърднетите в MCR имат добавени и деривационни релации (над 60 000 за език) както между различни части на речта (глагол – съществително, глагол – прилагателно и т.н.), така и между еднакви части на речта (две съществителни). На съществуващите деривационно свързани двойки не са приписани семантични релации.

Предложените морфосемантични релации в PWN имат няколко недостатъка:

(а) списъкът с релации не е пълен, т.е. съществуват деривационно свързани двойки, на които не съответства никоя от възможните релации;

(б) релациите са с неравномерен обхват – някои са много конкретни (например *агент*), докато други са значително по-общи и обхватни (например *събитие*);

(в) има релации, които могат да се припокриват (например *събитие* и *резултат*);

(г) не са разработени релациите между други части на речта, освен между съществителни и глаголи.

Заимстването на семантични релации от други уърднети също е проблематично поради следните причини:

(а) уърднетите ползват различен набор от релации, чието пренасяне в уърднетите за други езици често е неприложимо;

(б) потенциалните релации, заети от други уърднети, също са с неравномерен обхват – например за румънски има въведена релация *професия* в рамките на релацията *агент*, както и обща релация *свързан* (RELATED) без семантична конкретизация;

(в) уърднетите са с различни формати и различни условия за ползване на данните.

Значителна работа е проведена и върху морфосемантичните релации в БулНет – както пренасянето на морфосемантичните релации от PWN (Koeva 2008; Koeva et al. 2008), така и откриване на деривационни релации в български език (Dimitrova et al. 2014) и семантичната им класификация (Leseva et al. 2015).

4.2. Семантични релации

Някои части на речта, като например прилагателните и наречията, често се пренебрегват и не разполагат с добре дефинирана система от семантични релации, а остават под формата на изолирани синсети или групи от синсети. В последно време има направени някои опити за въвеждане на лексикално-семантични релации между прилагателните, мотивирана от лингвистични и прагматични съображения (Maziarz et al. 2012). Работи се и върху система от релации между прилагателните в български (Dimitrova and Stefanova 2016; Стефанова и Димитрова/Stefanova and Dimitrova 2016). Интересен би бил резултатът при добавяне на адвербиалните релации, представени от теорията на лексикалните функции (Mel'čuk 1996).

А. Хендлър (Handler 2007) представя метод за откриване на близки по значение думи с помощта на Word2Vec чрез представяне на думите във векторно пространство. Методът разпознава синонимия, хипонимия и хиперонимия, както и меронимия и холонимия. Може да бъде приложен за разширяване на БулНет с нови синсети и откриване на тяхното място в системата на Уърднет.

Някои автори (Siblini and Kosseim 2013), освен добре познатите семантични релации, разглеждат и релациите Gloss между даден синсет и пълнозначните думи (и съответните им синсети), които формират дефиницията на синсета. За целта се използва корпусът с глосове от PWN с отстра-

нена семантична многозначност. Този подход може да се пренесе и върху уърднети за други езици, включително БулНет, като се ползват както английските, така и българските дефиниции. Релациите са асоциирани с тежести въз основа на степента на близост, които могат да се използват за филтриране на релациите по тип.

4.3. Прагматични релации и други релации, извлечени от общи бази от знания

Различни изследователи започват да обръщат внимание на възможностите за обогатяване на Уърднет с допълнителна информация и релации, извлечени от енциклопедични бази от данни, общи бази от знания и др. Правят се опити за извличане на структурирани бази данни и онтологии от Wikipedia, DB-pedia, Yago, Freebase, BabelNet. Всеки от тези ресурси се отличават с добра структура и система от категории, което позволява сравнително лесното извличане на структурирана информация. Към момента преобладава използването им за извличане на именувани същности (топоними, имена на хора, имена на организации и др.), които могат да се добавят с релацията INSTANCE към съответните категории/синсети (Arnold and Rahm 2014).

Също така се използва извличането на терминология за специфични тематични области от съответните енциклопедични и други ресурси. Прилагат се методи за извличане и добавяне на растителни и животински видове, химични елементи, географски понятия и др.

Енциклопедичните ресурси могат да се използват и за приписване на релацията DOMAIN, която причислява даден синсет към специфична тематична област. Могат да бъдат използвани и класификациите на тематичните области. Също така могат да бъдат анализирани релациите между отделните елементи в базите данни (например връзките между статиите в Уикипедия), които да бъдат пренесени в Уърднет (например с помощта на открити синтактични шаблони и др.).

Семантични релации могат да се извличат и от различни корпуси. Корпусните данни се използват, като за всяко съществително се извлича и ранкира (по честота) групата от глаголи, които се срещат със съществителното (Yamada and Baldwin 2004). Високият ранк на определени глаголи свидетелства за вероятна семантична релация, свързана с реализацията на съществителното като аргумент с определена семантична роля в аргументната структура на глагола. Тези релации наподобяват морфосемантичните релации, но докато там връзката се установява чрез деривация, в случая се разчита на корпусни данни за установяване на семантична близост между лексикалните единици в речта.

4.4. Релации между Уърднет, ВърбНет, ФреймНет

За целите на синтактичния и семантичния анализ в приложения за машинен превод, отговаряне на въпроси и др. се прилагат комплексни зна-

ния от различни ресурси. Тъй като различните ресурси представят различна информация по специфичен начин, се правят опити за унифицирано представяне на данните чрез въвеждане на релации между единиците от различни ресурси като Уърднет, ВърбНет и ФреймНет (Shi and Mihalcea 2005).

ФреймНет (Baker et al. 2003; Коева/Коева 2010) предлага информация за синтактичната реализация на семантичните фреймове на предикатите и приписването на семантични роли на аргументите. ВърбНет (Kirpner et al. 2000), от друга страна, покрива синтактично-семантичните релации в поексплицитен вид, като дефинира селективните ограничения за аргументите в синтактичния фрейм. Уърднет от своя страна има много добро покритие на глаголната система както в PWN, така и в БулНет, като отразява и семантичните отношения на даден глаголен синсет с другите глаголи, но представя само ограничена и повърхностна информация за аргументната структура на глаголите.

Правят се опити да се конструира унифицирана рамка за комплексно представяне на информацията от трите ресурса (Shi and Mihalcea 2005) с цел да се постигне по-богата система за качествен и ефикасен анализ, която: (1) разширява покритието на ФреймНет, (2) ВърбНет се обогатява със семантична информация и (3) информация от Уърднет може да се използва за дефиниране на селективни ограничения.

Пациенца и др. (Pazienza et al. 2006) предлагат метод за интегриране на междуглаголните релации на Уърднет и семантичните фреймове на глаголите от ВърбНет с корпуса PennTreeBank с помощта на базата данни Prop Bank, която да представлява мост между езиковите ресурси и корпуса. Целта е представянето на едно по-пълно описание на семантиката на глаголите от междуглаголни релации към синтактични характеристики и от там – към повърхнинна реализация.

Комбинирането на ресурсите постига в цялост репрезентативните равнища на Пустейовски (1995, цитирани у Коева/Коева 2010), които дефинират семантичните граници, подходящи за употребата на дадено значение: лексикално-типизираща структура (отговаряща на Уърднет); аргументна и събитийта структура (отговарящи на информацията във ФреймНет и ВърбНет) и qualia структура – дефинира съществените свойства на обектите събитията и релациите, които се асоциират с лексикалната единица. Това може да доведе до формулиране на семантични правила за съотнасяне на различните семантични равнища с цел постигане на сполучлива семантична интерпретация на думата в контекст.

5. Заключение и насоки за бъдеща работа

Представеният анализ има за цел да очертае теоретичните и приложните аспекти на вече известни изследвания, фокусирани върху семантичните релации в различни уърднети, най-вече с оглед на възможностите за

обогатяване на БулНет с нови релации, както и за създаването на нови типове релации и релационни схеми.

Направихме обзор на теоретичните виждания за семантичните релации в Уърднет (и други подобни ресурси), които обуславят структурата на представяне на знанието, както и различните практически подходи към отразяването на семантичните релации между отделните единици. В този контекст са представени някои типове релации в чуждоезични ресурси и методите, използвани за тяхното извличане, които могат да бъдат приложени и за български език. В това число се включват възможности за разширяване на системата от морфосемантични релации с нови релации (например между съществителни), въвеждане на по-fino разделение на някои категории, идентифициране на релации от Уикипедия и други енциклопедични ресурси, интегриране на Уърднет с ресурси като ФреймНет, ВърбНет и др.

Семантичните мрежи от комплексно знание имат редица приложения в областта на компютърната лингвистика и обработката на естествен език и тяхното качество и покритие са от ключово значение за развитието на семантичния анализ, извличането на информация, машинния превод, автоматичното отговаряне на въпроси и много други.

Благодарности

Изследването е извършено в рамките на проекта „Семантична мрежа с широк спектър от семантични релации“, подкрепен от Фонд „Научни изследвания“ по програма „Финансиране на фундаментални научни изследвания“, Договор № 10/3/2016.

ЛИТЕРАТУРА

- Аристотел 1992: *Аристотел*. Категории. София, Наука и изкуство.
- Влахова 2005: *Влахова, Р.* „Винителни“ и „дателни“ префигирани глаголи в съвременния български език. – В: Коева, Св. (съст.). Аргументна структура. Проблеми на простото и сложното изречение. София, СемаРШ, с. 80–105.
- Влахова 2006: *Влахова, Р.* Семантични и синтактични особености на префигираните глаголи с прономинален елемент в български. – В: Езиковедски изследвания в чест на 75-годишнината от рождението на ст.н.с. I ст. д.ф.н. Йордан Пенчев. София, „Артграф“, с. 320–329.
- Джонова 2005: *Джонова, М.* Информационна структура на изреченията с предикати, изискващи аргумент експериенцер. – В: Юбилеен славистичен сборник. Благоевград, Университетско издателство „Неофит Рилски“, с. 325–332.
- Джонова 2008: *Джонова, М.* Предикати със задължителен аргумент експериенцер от типа „дострашее ме“, „домъчнее ми“. В: Сборник в чест на проф. Василка Радева. София, 2008, с. 568–577.

- Ельмслев 1960: *Ельмслев, Л.* Прологомени к теории языка. (Пер. с англ.). – Новое в лингвистике. Вып. 1. Москва, 1960, 264–389.
- Коева 2006: *Коева, Св.* (съст.). Аргументна структура: Проблеми на простото и сложното изречение. (*Verba magistri*). София, СЕМА РШ.
- Коева 2007: *Коева, Св.* Релациите в Wordnet. – В: Лексикологията и лексикографията в съвременния свят. Велико Търново, „Знак’94“, с. 287–297.
- Коева 2008: *Коева, Св.* (съст.) Българският ФреймНет. Семантико-синтактичен речник на българския език. София, Издателство на БАН.
- Коева 2010: *Коева, Св.* Българският ФреймНет. София, Институт за български език при БАН.
- Коева 2014: *Коева, Св.* WordNet и БулНет – В: Езикови ресурси и технологии за български език. Академично издателство „Проф. Марин Дринов“, с. 154–173.
- Ницолова 2005: *Ницолова, Р.* За формалното означаване на семантичните роли в българските конструкции с отглаголни съществителни. – В: Коева, Св. (съст.). Аргументна структура. Проблеми на простото и сложното изречение. София, СемаРШ, с. 69–79.
- Пернишка 2014: *Пернишка, Ем.* Лексикална семантика. Структурно-семантични особености на българската лексика. – В: Крумова-Цветкова, Л., Благоева, Д. Колковска, С., Пернишка, Е., Божилова, М. Българска лексикология. Т. 1. София, цитати на стр. 425; 429; 433; 434.
- Стефанова и Димитрова 2016: *Стефанова, В., Цв. Димитрова.* Прилагателното име в Българския уърднет. – Български език, 4, 2016, с. 1–15.

REFERENCES

- Aristotel 1992: *Aristotel.* Kategorii [Categories]. Sofia, Nauka I izkustvo.
- Arnold and Rahm 2014: *Arnold, Patrick, and Erhard Rahm.* Extracting Semantic Concept Relations from Wikipedia. – In: *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS '14)*. ACM, New York, NY, USA, 2014.
- Baker et al. 2003: *Baker, C. F., Fillmore, C. J., and Cronin, B.* The Structure of the FrameNet Database. – In: *International Journal of Lexicography*, vol. 16, 281–296, 2003.
- Bilgin et al. 2004: *Bilgin, Orhan, Çetinoglu, Özlem, and Oflazer, Kemal.* Morphosemantic Relations In and Across Wordnets. A Study Based on Turkish. – In: *Proceedings of the Second International WordNet Conference (GWC 2004)*. <http://www.fi.muni.cz/gwc2004/proc/gwc2004-book#page=72>
- Bond and Foster 2013: *Bond, Francis, Foster, Ryan.* Linking and extending an open multilingual wordnet. – In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: ACL-2013*. Sofia, 1352–1362.
- Caplan and Herrmann 1993: *Caplan, L. J., Herrmann, D. J.* Semantic relations as graded concepts. – *Zeitschrift für Psychologie*, 201, 8–97.
- Chafe 1970: *Chafe, W. L.* Meaning and the Structure of Language. University Press, Chicago.
- Cruse 2002: *Cruse, D. A.* Lexical semantics. Cambridge, Cambridge University Press.

- Cruse 2011: *Cruse, D. A.* Meaning in language: An introduction to semantics and pragmatics (3rd ed.). Oxford, Oxford University Press.
- Dimitrova et al. 2014: *Dimitrova, Tsv., Ek. Tarpomanova, B. Rizov.* Coping with derivation in the Bulgarian WordNet. – In: *Proceedings of the Seventh Global Wordnet Conference (GWC 2014)*, 109–117.
- Dimitrova and Stefanova 2016: *Dimitrova, T., V. Stefanova.* Adjectives in WordNet: Semantic Issues. – In: *Proceedings of the 12th International Conference Linguistics Resources and Tools for Processing the Romanian Language (ConsILR-2016)*. Faculty for Computer Science, Alexandru Ioan Cuza University, Iași, 2016.
- Dzhonova 2005: *Dzhonova, M.* Informatsionna struktura na izrecheniyata s predikati, iziskvashti argument eksperientser. [Information Structure of Sentences with Predicates Requiring an Experientser Argument]. – In: *Yubileen slavistichen sbornik*, Blagoevgrad. “UI Neofit Rilski”, 325–332.
- Dzhonova 2008: *Dzhonova, M.* Predikati sas zadalzhitelten argument eksperientser ot tipa „dostrashee me“, „domachnee mi“ [Predicates with Obligatory Argument Experientser from the Type „dostrashee me“, „domachnee mi“]. – In: *Sbornik v chest na prof. Vasilka Radeva*. Sofia, 568–577.
- Fellbaum 1998: *Fellbaum, C.* (ed.). WordNet: An Electronic Lexical Database. Cambridge, MA, MIT Press.
- Fellbaum et al. 2009: *Fellbaum, Chr., A. Osherson, P. E. Clark.* Putting Semantics into WordNet's „Morphosemantic“ Links. – In: *Human Language Technology. Challenges of the Information Society*, Zygmunt Vetulani and Hans Uszkoreit (Eds.). Lecture Notes In Artificial Intelligence, Vol. 5603. Springer-Verlag, Berlin, Heidelberg, 350–358.
- Fillmore 1982: *Fillmore, C. J.* Frame Semantics. – In: *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea, Soeul, Hanshin, 111–137.
- Goldberg 1994: *Goldberg, A.* Constructions: A Construction Grammar Approach to Argument Structure. University of Chicago Press.
- Gonzalez-Agirre et al. 2012: *Gonzalez-Agirre, A., E., Laparra and G. Rigau.* Multilingual Central Repository version 3.0. 8th international conference on Language Resources and Evaluation (LREC'12). Istanbul, Turkey, 2012.
- Green 2001: *Green, R.* Relationships in the organization of knowledge: An overview. – In: *C.A. Bean & R. Green* (Eds.), Relationships in the organization of knowledge. Dordrecht, Kluwer, 3–18.
- Green 2002: *Green, R.* Internally-structured conceptual models in cognitive semantics. – In: *R. Green, C. Bean, & S. H. Myaeng.* The semantics of relationships: An interdisciplinary perspective. Dordrecht, Kluwer, 73–90.
- Handler 2007: *Handler, Abram.* An empirical study of semantic similarity in WordNet and Word2Vec. A Thesis Submitted to University of New Orleans, Columbia University, 2007. <http://scholarworks.uno.edu/cgi/viewcontent.cgi?article=3003&context=td>
- Hjelmsiev 1960: *Hjelmsiev, L.* Prolegomeny k teorii yazyka. [Prolegomena to the Theory of Language]. – *Novoe v lingvistike*. Moskva, 1960, 264–389.
- Jackendoff 1990: *Jackendoff, R. S.* Semantic Structures. Cambridge, Mass., The MIT Press, 1990.

- Jakobson and Halle 1956: *Jakobson, R., M. Halle*. Fundamentals of language. The Hague, Mouton.
- Khoo and Na 2007: *Khoo, Christopher, Jin-Cheon Na*. Semantic relations in information science. – Annual Review of Information Science and Technology. Volume 40, Issue 1, January 2007, 157–228.
- Kipper et al. 2000: *Kipper, Karin, Hoa Trang Dang and Martha Palmer*. Class-based construction of a verb lexicon. – In: *Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-2000)*. Austin, TX, July-August, 2000, 691–696.
- Koeva 2006: *Koeva, S.* Argumentna struktura: Problemi na prostoto i slozhnoto izrechenie [Argument Structure: Problems of the Clause and of the Complex sentence]. (Verba magistri). Sofia, SEMA RSH.
- Koeva 2007: *Koeva, Sv.* Relatsiite v WordNet [The Relations in WordNet]. – In: *Leksikologiyata i leksikografiyata v savremenniya svyat*. Veliko Tarnovo. “Znak’94”, 287–297.
- Koeva 2008a: *Koeva, Sv.* (ed.). Balgarskiyat Freymnet. Semantiko-sintaktichen rechnik na balgarskia ezik [The Bulgarian FrameNet. A Semantic-syntactic Dictionary of the Bulgarian Language]. Sofia, Izdatelstvo na BAN.
- Koeva 2008b: *Koeva, Sv.* Derivational and Morphosemantic Relations in Bulgarian Wordnet. – In: *Intelligent Information Systems. XVI*, Warsaw, Academic Publishing House, 359–389, 2008.
- Koeva et al. 2008: *Koeva, Sv., Cv. Krstev, D. Vitas*. Morpho-semantic Relations in Wordnet – a Case Study for two Slavic Languages. – In: *Tanács, Attila; Csendes, Dora; Vincze, Veronika; Fellbaum, Christiane & Vossen, Pietr*. Global Wordnet Conference, 2008, Szeged, Hungary, University of Szeged, Department of Informatics, 239–253, Global Wordnet Conference (GWC’08). <http://poincare.matf.bg.ac.rs/~cvetana/biblio/Krstev-GWN08.pdf>
- Koeva 2010: *Koeva, Sv.* Balgarskiyat Freymnet [The Bulgarian FrameNet]. Sofia, Institut za balgarski ezik.
- Koeva 2014: *Koeva, Sv.* WordNet I Bulnet [WordNet and BulNet]. – In: *Ezikovi resursi i tehnologii za balgarski ezik*. Sofia. Akademichno izdatelstvo “Prof. Marin Drinov”, 154–173.
- Koeva et al. 2016: *Koeva, S., S. Leseva, I. Stoyanova, T. Dimitrova, M. Todorova*. Automatic Prediction of Morphosemantic Relations. – In: *Proceedings of the Eighth Global Wordnet Conference*, 168–176. <http://gwc2016.racai.ro/proceedings.html>
- Langacker 1987: *Langacker, R.* Nouns and Verbs. – In: *Language*, Vol. 63, Issue 1 (March, 1987), 53–94. <http://terpconnect.umd.edu/~israel/Langacker-NounsVerbs.pdf>
- Leseva et al. 2015: *Leseva, Sv., M. Todorova, Ts. Dimitrova, B. Rizov, Iv. Stoyanova and Sv. Koeva*. Automatic Classification of WordNet Morphosemantic Relations. – In: *Proceeding of the 5th Workshop on Balto-Slavic Natural Language Processing (BSNLP 2015)*, Hissar, Bulgaria, 2015.
- Levin 1993: *Levin, B.* English Verb Classes and Alternations: A Preliminary Investigation. Chicago, The University of Chicago Press.

- Longacre 1976: *Longacre, R. E.* An Anatomy of Speech Notions. Peter de Ridder Press.
- Lyons 1977: *Lyons, J.* Semantics (2 vols.). Cambridge, Cambridge University Press.
- Lyons 1995: *Lyons, J.* Linguistic semantics: An introduction. Cambridge, Cambridge University Press.
- Maziarz et al. 2012: *Maziarz, Marek, Stanisław Szpakowicz, Maciej Piasecki.* Semantic relations among adjectives in Polish Wordnet 2.0: a new relation set, discussion and evaluation. – *Cognitive Studies | Études cognitives*, vol. 12, 2012, 149–180.
- Mel'čuk 1996: *Mel'čuk, I.* Lexical functions: A tool for the description of lexical relations in a lexicon. – In: *L. Wanner* (ed.). Lexical functions in lexicography and natural language processing. Amsterdam, John Benjamins, 37–102.
- Mel'čuk and Žolkovskij 1970: *Mel'čuk, I., A. Žolkovskij.* Towards a functioning meaning-text model of language. – *Linguistics*. 1970, 8(57), 10–47.
- Miller 1995: *Miller, G. A.* WordNet: A Lexical Database for English. – In: *Communications of the ACM*. Vol. 38, No. 11, 39–41.
- Mititelu 2012: *Mititelu, Verginica Barbu.* Adding Morpho-semantic Relations to the Romanian Wordnet. – In: *Proceedings of LREC*, 2012.
- Murphy 2003: *Murphy, M. L.* Semantic relations and the lexicon: Antonym, synonymy, and other paradigms. Cambridge, Cambridge University Press.
- Nitsolova 2005: *Nitsolova, R.* Za formalnoto oznachavane na semantichnite roli v balgarskite konstruksii s otglagolni sashtestvitelni [For the Formal Marking of the Semantic Roles in the Bulgarian Constructions with Verbal Nouns. – In: *Koeva, Sv* (ed.). Argumentna struktura. Problemi na prostoto i slozhnoto izrechenie. Sofia, “SemaRSH”, 69–79.
- Pala and Hlavackova 2007: *Pala, K., D. Hlavackova.* Derivational relations in Czech Wordnet. – In: *Proceedings of the Workshop on Balto-Slavonic Natural Language Processing*, 2007, 75–81.
- Pazienza et al. 2006: *Pazienza, Maria Teresa, Pennacchiotti, Marco, Zanzotto, Fabio Massimo.* Mixing WordNet, VerbNet and PropBank for studying verb relations. – In: *Proceedings of the 5th Language Resource and Evaluation Conference (LREC 2006)*. Genova, Italy, 2006.
- Pernishka 2014: *Pernishka, E.* Leksikalna semantika. Strukturno-semantichni osobnosti na balgarskata leksika [Lexical Semantics. Structural and Semantic Properties of the Bulgarian Lexis]. – In: *Krumova-Tsvetkova, L., D. Blagoeva, S. Kolkovska, E. Pernishka, M. Bozhilova.* Bulgarska lexicologia Vol. 1. Sofia.
- Piasecki and Koeva 2017: *Piasecki, M., Koeva, S.* WordNet Relations in the Bulgarian-Polish Bilingual Perspective. – In: *Доклади от Международната юбилейна конференция на Института за български език „Проф. Любомир Андрейчин“* (София, 15–16 май 2017 година). София, ИБЕ, 2017, с. 285–297. http://ibl.bas.bg/wp-content/uploads/2017/05/SBORNIK_IBE_15-16May2017.pdf
- Pustejovsky 1995, 1998: *Pustejovsky, J.* The Generative Lexicon. The MIT Press. Cambridge, Massachusetts, London, England.

- Sapir 1921: *Sapir, E.* Language: An introduction to the study of speech. New York, Harcourt, Brace and Company.
- Shi and Mihalcea 2005: *Shi, L., R. Mihalcea.* Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing. – In: *Gelbukh A.* (eds) Computational Linguistics and Intelligent Text Processing. CILing 2005. Lecture Notes in Computer Science, vol 3406. Springer, Berlin, Heidelberg.
- Siblini and Kosseim 2013: *Siblini, R., L. Kosseim.* Using a Weighted Semantic Network for Lexical Semantic Relatedness. – In: *Proceedings of Recent Advances in Natural Language Processing.* Hissar, Bulgaria, 7–13 September 2013, 610–618.
- Spellman et al. 2001: *Spellman, B. A., K. Holyoak & R. G. Morrison.* Analogical priming via semantic relations. – In: *Memory & Cognition*, 29 (3), 383–393.
- Stefanova and Dimitrova 2016: *Stefanova, V., Ts. Dimitrova.* Prilagatelnoto ime v Balgarskiya uardnet [The Adjective in the Bulgarian WordNet]. – *Balgarski ezik*, 2016, № 4.
- Thellefsen 2010: *Thellefsen, M.* Knowledge Organization, Concepts, Signs: A Semiotic Framework. – In: *Royal School of Library and Information Science.* Aalborg, 2010. xi, 193 p. http://pure.db.dk/da/publications/knowledge-organization-conceptssigns_cache198-b26c-4987-abcc-e33a0a8ed910.html
- Tufiş et al. 2004: *Tufiş, Dan, Dan Cristea and Sofia Stamou.* BalkaNet: Aims, Methods, Results and Perspectives. A General Overview. – In: *D. Tufiş* (ed.): Special Issue on BalkaNet. Romanian Journal on Science and Technology of Information, 7 (3–4), 9–44.
- Van Valin and LaPolla 1997: *Van Valin Jr., R. D., LaPolla, R. J.* Syntax: Structure, meaning and function. Cambridge University Press.
- Vlahova 2005: *Vlahova, R.* „Vinitelni“ i „datelni“ prefigirani glagoli v savremeniya balgarski ezik. [Genitive and Dative Prefixed Verbs in the Contemporary Bulgarian Language]. – In: *Koeva, Sv.* (ed.). Argumentna struktura. Problemi na prostoto i slozhnoto izrechenie. Sofia, SemaRSH, 80–105.
- Vlahova 2006: *Vlahova, R.* Semantichni i sintaktichni osobenosti na prefigiranite glagoli s pronominalen element v balgarski [Semantic and Syntactic Properties of the Prefixed Verbs with Pronominal Element in Bulgarian]. – In: *Ezikovedski izsledvaniya v chest na 75-godishninata ot rozhdenieto na st.n.s. I st. d.f.n. Yordan Penchev.* Sofia, „Artgraf“, 320–329.]
- Vossen 1998: *Vossen, Piek* (ed.). Euro WordNet. Kluwer.
- Vossen 2005: *Vossen, Piek.* Building wordnets. <http://www.globalwordnet.org/gwa/BuildingWordnets.ppt>.
- Wellisch 2000: *Wellisch, H. H.* Glossary of terminology in abstracting, classification, indexing, and thesaurus construction. 2nd ed. Medford. – In: *Information Today, Inc.*
- Yamada and Baldwin 2004: *Yamada, Ichiro, Timothy Baldwin.* Automatic Discovery of Telic and Agentive Roles from Corpus Data. – In: *Proceedings of PACLIC 18*, December 8th–10th, 2004, Waseda University, Tokyo, 115–126.

РЕЗЮМЕ

Статията представя преглед на популярни теоретични описания на семантични релации, както и приложните аспекти на системите от семантични релации в уърднети на различни езици с оглед на възможностите за разширяването на Българската лексикално-семантична мрежа БулНет с нови семантични релации.

Ключови думи: семантични релации, Уърднет, БулНет

✉ *Д-р Ивелина Стоянова, гл. ас. д-р Мария Тодорова*

Секция по компютърна лингвистика

Институт за български език „Проф. Л. Андрейчин“ при БАН

бул. „Шипченски проход“ 52, бл. 17, 1113 София, България

✉ *Ivelina Stoyanova, PhD, Assist. Prof. Maria Todorova, PhD*

Department of Computational Linguistics

Institute for Bulgarian Language, Bulgarian Academy of Sciences

52 Shipchenski prohod, Bl. 17, 1113 Sofia, Bulgaria