

**ОСОБЕНОСТИ НА СЕМАНТИЧНАТА И СИНТАКТИЧНАТА
СТРУКТУРА НА СОБСТВЕНИТЕ ИМЕНА ЗА ЛИЦА,
ГЕОГРАФСКИ ОБЕКТИ И ОРГАНИЗАЦИИ
В БЪЛГАРСКИ ЕЗИК**

Светла Коева, Цветана Димитрова
Институт за български език, БАН

**ON SEMANTIC AND SYNTACTIC STRUCTURE OF PERSON,
LOCATION AND ORGANIZATION NAMED ENTITIES
IN BULGARIAN**

Svetla Koeva, Tsvetana Dimitrova
Institute for Bulgarian Language
Bulgarian Academy of Sciences

Резюме. Изследването, представено в настоящата студия, предлага единен модел за описание на три основни типа съществителни собствени имена – названия на лица, географски обекти и организации. Предлага се структурна и семантична класификация на разглежданите съществителни собствени имена и модел за езиково независимо семантично описание. Представени са основните типове семантични модели, характерни за названията на лица, географски обекти и организации, и общо описание на синтактичните структури, които им съответстват в български.

Ключови думи: собствени имена, семантика, синтаксис

Abstract. The present study offers a common model for description of three classes of named entities in Bulgarian – persons, locations, and organizations – which can be expressed in single words, multiword expressions, and noun phrases. We offer a structural and semantic classification of the named entities and a language-independent model for description of the semantic and syntactic structure of proper names according to their class. Further, we present the major types of semantic structures for persons, locations, and organizations and a general description of the corresponding syntactic structures in Bulgarian.

Keywords: named entities, semantics, syntax

1. Въведение

Настоящата студия разглежда названията на лица (*Иван, Иван Петров*), географски обекти (*Апенини, Черно море*) и организации (*„Гугъл“, Организация на обединените нации*), както и словосъчетанията с тях (*къдравата Ани, професорът по физика Валери Димитров, новата фирма за дограма „Саламандър“, обединеният Европейски съюз, пълноводната река Марица, обновеният град Стара Загора*). Изследването предлага: а) единен модел за описание на именуваните обекти (общо название за имена на лица, имена на организации, географски названия – думи, съставни думи и словосъчетания с тях); б) класификация на именуваните обекти, съобразена със съществуващите класификации, но мотивирана от семантичната и синтактичната структура на имената, обединени в отделен клас; в) описание на типовете семантични модели, с които се характеризират класовете имена на лица, организации и географски обекти; г) обща характеристика на синтактичните структури, които кореспондират на семантичните модели за български.

Резултатите от представеното изследване може да послужат при формулирането и приложението на методи за разпознаване на именуваните обекти в задачи от областта на компютърната обработка на естествения език, включително за: определяне на най-важните фрази в даден текст; категоризация и клъстеризация на документи в определени тематични области; определяне на документи, близки по съдържание; доказване на авторство на даден документ; автоматично извличане на кратко резюме на текст и др. Резултатите от анализа на структурите могат да се използват при разработването на лингвистични ресурси като: речници (спрямо дефинирани класификационни категории за именуваните обекти, съдържащи собствени имена, географски названия, имена на организации и информация за тяхното формообразуване); правила за разпознаване на различни категории и подкатегории именувани обекти; анотирани корпуси за тестване на резултатите от приложението на правилата.

Предлаганото от нас описание на именуваните обекти в качеството им на лингвистични обекти минава през няколко равнища – семантично (структурни модели), синтактично (структурни модели: конституентност, словоред, съгласуване, обособяване, алтернации), морфологично (части на речта, граматични характеристики, използвани при съгласуване) и графично (главни букви, абревиатури, пунктуация). Семантичното равнище е до голяма степен езиково независимо, докато останалите са езиково зависими.

В следващия раздел от студията ще представим накратко някои възгледи за същността на именуваните обекти и техните характеристики както с оглед на конституентността на фразата, така и по отношение на референтността им, която ще бъде по-подробно засегната в раздела, представящ различни класификации на именуваните обекти, както и класификацията, която се предлага в нашето изследване. Представят се някои от езиковите ресурси, които бяха

подготвени в хода на изследването и върху които бяха направени и някои от наблюденията. По-нататък се анализират по-подробно семантичните модели на трите класа (и техните подкласове) именувани обекти и се предлага обща характеристика на синтактичните структури, чрез които се изразяват тези модели.

2. Именувани обекти

Именуваните обекти (или наименования) са елементи, които са важни за разбирането на текста, тъй като посочват кой, какво, на кого, къде и кога казва нещо в даден текст, както и на кого, къде и кога се случват описаните в текста събития. Като именувани обекти най-често се категоризират имена на лица (*Росен, Росен Иванов Плевнелиев, Луи XIV, Любен Дилов-син, Фон Мюнхаузен, Бан Ки-мун, Пиер Огюст Реноар, Имануел Кант, Пипи Дългото Чоранче*); имена на организации (*Организацията на обединените нации, ООН, Министерство на културата, Институт за български език, Пощенска банка, „Майкрософт“, „Морган Стенли“*); географски понятия (*София, Атли, Средиземно море, булевард „Шипченски проход“*) – тези три класа са известни под наименованието *epitex* (Надо, Секине/Nadeau, Sekine 2007). В други класификации като именувани обекти се категоризират и числови изрази – за време (*24 май 2013 г., септември, 12 часа*), за количество (*30 г, 30 грама*), за стойност (*\$20 000, \$20 хил.*), термини (наименования на гени, химични елементи, физични величини); наименования на произведения на изкуството (картини, скулптури, филми, литературни произведения); наименования на лекарства; наименования на продукти на материалната култура (марки, напитки, хранителни продукти) и др. (Надо, Секине/Nadeau, Sekine 2007; Секине, Нобата/Sekine, Nobata 2004).

Според Крипке (Kripke 1980) наименованията имат способността да означават референта точно, еднозначно и непротиворечиво – като един и същи обект във всеки възможен свят. Това означава, че наименованията са еднозначни означаващи, рефериращи към извънезикови (единични и уникални) обекти. В теорията на т.нар. строги определители (*rigid designators*, вж. Крипке/Kripke 1980; Пътнам/Putnam 1975) референцията се осъществява чрез връзка с обекта, прикрепването на наименованието към обекта често е случайно.

Съществителните собствени имена сами по себе си не описват характеристиките на обекта, към който реферират, но са логически свързани с него и наследяват неговите характеристики (Сърл/Searle 1958, 173), т.е. ако *Елена* е град, може да е възрожденски, китен или старопланински, но ако *Елена* е жена, може да е висока, руса или къдрава. Макар че собствените имена служат за уникално назоваване на даден обект, с което той става различим в множество от еднородни обекти (например конкретната жена с име Елена е различима в множеството от всички жени), принадлежността на обектите

към дадено множество с конкретни свойства предопределя възможността на семантично равнище да се обособят семантични структури, които описват свойствата на множеството (класа). Структурата на фразите с именуваните обекти (предимно именни фрази), както и възможните зависими конститuentи (включително съществително и именна фраза – *Дирекция „Научни изследвания“*; прилагателно – *Международен валутен фонд*; предложна фраза – *Организация на обединените нации*) могат да се опишат на синтактично равнище. В настоящото изследване разглеждаме бащините и фамилните имена като прилагателни за притежание, съответно *Иван Петров* е словосъчетание с главна част лично име; ако се употребяват самостоятелно, фамилните имена се разглеждат като субстантивирани прилагателни (Коева, Димитрова/Коева, Dimitrova 2015). В структурата на словосъчетания с главна част лично име, географско название или име на организация се срещат прилагателни и словосъчетания с главна част прилагателно, съществителни и словосъчетания с тях, които са приложения, както и предложни групи. Подчинените определителни изречения към собствено име са извън обекта на настоящото изследване. Разглеждаме обаче допустимите показателни и притежателни местоимения, позицията на притежателните клитики и на въпросителните клитики.

Описанието на именуваните обекти включва морфологични характеристики – например част на речта на съставните думи: в състава на именуваните обекти по-често се срещат съществителни и прилагателни (*Съединени американски щати, Българска народна банка*), включително предложни фрази (*Комисия за финансов надзор, Организация на обединените нации*). Съществителните собствени имена за лица са в женски и мъжки род, а за организации и географски обекти – и в трите рода, включително плуралия тантум. И трите вида имена могат да бъдат както прости думи, така и сложни думи (*Мария, Димитровград, „Данон“*) или съставни думи (*Йоан Павел, Стара планина, България Еър*).

Описанието на именуваните обекти на графично равнище включва характеристиките на поредиците от символи, които участват в наименованието: дали думата започва с главна буква; дали е част от поредица от думи, започващи с главни букви (*Антон Петров Иванов, Лос Анджелис, Саудитска Арабия*); дали поредицата от символи (думата) е изписана само с главни букви (в случаите на лексикални съкращения с поредица от главни букви – *ООН, САЩ, ЮАР, КФН*); дали използваната графична система се различава (например в текстове на кирилица често корпоративните наименования се срещат изписани на латиница – *Lehman Brothers, Apple*); дали съдържа римска цифра (*Йоан Павел II*); каква пунктуация е използвана: дефис (при *Любен Дилевсин, Бан Ки-мун, Мохамед ал-Барадей* или *Мохамед ел-Барадей*), поредица от буквени символи, завършващи с точка (например при съкращенията – *Ал. Ив. Петров*), кавички (*Софийски университет „Св. Климент Охридски“*). Тези характеристики са езиково специфични (например в немски главните бук-

ви, с които започват всички съществителни имена (наименованията често са именни фрази), няма да са от такова значение, както в езици като английски, български, френски).

3. Класификация

Една от задачите на представеното изследване е свързана с класифицирането на именуваните обекти – имена на лица, на географски обекти и на организации – и описанието на семантичната и синтактичната структура на отделните класове.

Хъдълстън (Huddleston 1988, 96) посочва, че в английски терминът „съществително собствено“ (proper noun) се използва за едносъставни имена (*Иван, Франция, „Кауфланд“*), а терминът „собствено име“ (proper name) се отнася до функцията на именуваните обекти да реферират към определени обекти от действителността, като тази функция може да се изпълнява както от отделни думи, така и от съставни лексикални единици (*Иван Иванов, Доналд Тръмп-младши, Съединени американски щати, „Лимън Брадърс“, Министертво на правосъдието*).

За целите на разпознаването и най-вече на класифицирането на именуваните обекти и особено при липса на предварителна информация за категориите (например *Джон Диъри* е името на основателя на американската компания за селскостопанска техника „*Диъри & Ко.*“, чиято търговска марка обаче също е „*Джон Диъри*“) често се изследва непосредственият контекст, в който се намират думи или фрази, назоваващи общия клас обекти от действителността, към които се причислява референтът на наименованието (*компания, президент, директор, генерал, професор, проф. д-р*). Лингвистичните методи за описание и класификация на именуваните обекти разчитат на лингвистичен анализ, за да опишат лексикалния състав и характеристики на елементите, от които се състоят именуваните обекти, като на тази основа се формулират правила с включени семантични и лексикално-синтактични ограничения.

Именуваните обекти могат да бъдат класифицирани на различен принцип, като най-общата и най-често използваната класификация ги разделя на три основни категории, а именно – наименования на лица, наименования на организации и наименования на географски обекти, като към тези три категории често се добавят и числовите изрази (Надо, Секине/Nadeau, Sekine 2007). Една от най-подробните класификации на именувани обекти включва над 200 категории (Секине, Нобата/Sekine, Nobata 2004), сред които наименования на: болести, минерали, теории, конференции, влакове, престъпления и др. Например подкласовете за наименованията на лица, използвани от Флайшман и Хови (Fleischman, Нову 2002) и Джулиано (Giuliano 2009), включват: създател (*автор – музикант, художник, режисьор*), комуникатор (*представител, писател – поет, драматург*), учен (*физик, химик, социолог, математик,*

биолог), медицинско лице, човек, който се занимава с бизнес, човек, който изнася представления (*музикант, актьор*). Географските наименования пък се разпределят в подкласове на наименования на: летища, държави, градове, континенти, вулкани, върхове, пустини, гори, заливи, езера, острови, реки и др. (Ферес и др./Ferrés et al. 2004). Числовите изрази се разделят на изрази за време, за количество, за валута, за измерване (в различни области) (Надо, Секине/Nadeau, Sekine 2007). В литературата са описани различни йерархии, тъй като практически всяко съществително нарицателно, което може да определя референта на съществителното собствено, може да го характеризира по специфичен начин.

За идентификация и класификация на именуваните обекти най-общо се използват два вида контекстови „свидетелства“ – външни и вътрешни (*external evidences – internal evidences*) (МакДоналд/McDonald 1996). Вътрешните свидетелства са елементи, които са включени в името, а външните свидетелства се добиват от контекста, в който е употребено името. Например вътрешни могат да бъдат елементи, които показват какъв е типът на дадена компания – дали правната форма, под която е регистрирана, е *ООД, ЕООД, АД*. Външните свидетелства за референта на името са свързани с функцията на собствените имена да назовават един от множество обекти, принадлежащи към специфичен тип (*професор, банка, река*). Наличието на външните и вътрешните свидетелства в непосредствения контекст на появата на собственото име в даден текст осигурява условия, при които собственото име може да се категоризира по определен начин.

Други автори (Манини и др./Magnini et al. 2002) разглеждат думите, които категоризират именуваните обекти, в йерархията на лексикално-семантичната мрежа Уърднет, в която са включени синонимни множества, назоваващи категориите, към които могат да принадлежат именуваните обекти. Съставени са големи групи от съществителни, които могат да се използват като характеризиращи референта на дадено собствено име – хипонимите на редица синсети в най-горната част на хиперонимно-хипонимното дърво като {човек; лице}, {място} и {организация}. Например за лица те използват 6775 класа и с помощта на правила разпознават и класифицират собствените имена в различни конструкции. Връзката между съществителните собствени и понятието, което категоризира техния референт, в Уърднет е наречена хипероним на собствено име – хипоним на собствено име (*instance hypernym – instance hyponym*) (Родригес и др./Rodríguez et al. 1998), като собствените имена наследяват характеристики от йерархията, към която принадлежат.

Друга позната класификация е тази на Андерсън (Anderson 2007), която подробно класифицира собствените имена, както следва: прости непрозрачни собствени имена (*Елена*); прости собствени имена, които произхождат от нарицателни имена (*Радост*); собствени имена, свързани с други собствени имена („*Линкълн*“ – име на булевард, който носи името на личност); собствени имена, които са деривативно производни от други собствени имена (сла-

вянските фамилни имена); сложни собствени имена, някои от които съдържат друго име (*Куинсланд, Нютаун*); съставни собствени имена, някои от които съдържат друго име (*Университетът на Куинсланд*); собствени имена, които са цели фрази или изречения (например заглавия на книги, филми и под.).

Една от класификациите на именуваните обекти в български – на Лесева и Стоянова (Лесева, Стоянова/Leseva, Stoyanova 2007) – ги категоризира в три класа: същински наименования, дескриптивни наименования и каскадни наименования. Същинските наименования се отнасят към конкретен, еднозначно определен референт – най-често това са същински собствени имена (*Джон Мейджър, Организация на обединените нации, Черно море, „Гугъл“* и др.). Дескриптивните наименования включват т.нар. дескриптори, без собствени имена (*президентът на Република България, главният изпълнителен директор на „Гугъл“*). Каскадните наименования се състоят от дескриптор и същинско име (*президентът Росен Плевнелиев, прибалтийската република Латвия*) (Стоянова/Stoyanova 2014).

В своето изследване върху съставните лексикални единици Стоянова (Stoyanova 2014) разглежда структурата на именуваните обекти, като използва основните категории дескриптори и дескриптивни наименования по класификацията на „Би Би Ен Текнолоджис“ (BBN) и различни видове именуван обекти (Брунщайн/Brunstein 2002), като ги обобщава и допълва според особеностите на българския език.

Класификацията на съставните наименования на Тодорова (Todorova 2010) разделя имената на: антропоними (включително поликомпонентни лични имена – *Анна-Мария*, прякори – *Поразяващата уста*, псевдоними – *Елин Пелин*); топоними (*Велико Търново, Луда Камчия*), космоними (*Млечен път, Слънчева система*), хромоними (*Втората световна война, Ден на славянската писменост и култура*), хрематоними – наименования на предмети от материалната култура, имена на картини, статуи, произведения на изкуството, търговски марки – *„Кока Кола“*, *„Видима идеал стандарт“*; ергоними – имена на учреждения, организации, институции, компании и под.

Тъй като нашето изследване е насочено към определяне на структурата и конституентността на именната фраза, която е именуван обект, изготвихме две класификации: едната отразява структурата и границите на именуваните обекти (структурна класификация), а втората – семантичните класове на именуваните обекти, които предопределят семантичната и синтактичната им структура (семантична класификация). Имената на лица, имената на географски обекти и имената на организации могат да се изразяват както от отделни думи – съществителни собствени (*Иван, „Епъл“*), така и от съставни лексикални единици (*Иван Иванов, Българска асоциация за семейно планиране*). В структурната класификация отразяваме: 1) дали името е съставна лексикална единица, или е отделна дума – съществително собствено; 2) дали съставната лексикална единица задължително включва и категоризиращо нарицателно име (*Международен валутен фонд*); 3) дали наименованието може да бъде

специфицирано от категоризиращо нарицателно име в рамките на именно словосъчетание („УниКредит“ и банка „УниКредит“). Тъй като проблем при автоматичното разпознаване на именуваните обекти може да възникне при определянето на границите им в случаите, в които едно наименование се включва в състава на друго (в литературата на английски това са т.нар. nested named entities – например Фондация „Конрад Аденауер“, където именуваният обект включва лично име в състава на име на организация), тази възможност беше отразена в структурната класификация. Друг следван критерий е конституентността: отделните думи – съществителни собствени (*Мария, Европа, „Фейсбук“*) и съставни лексикални единици (*Мария Петрова, Южна Америка, „Хеброс Банк“*), могат да се разширяват в именни фрази (*професор Мария Петрова, град Лос Анджелис, Организация на обединените нации*).

Следвайки изведените по-горе принципи, формулирахме следната структурна класификация на именуваните обекти (лица, организации, географски обекти) със съответните подкласове.

3.1. Наименования на лица

3.1.1. Съществително собствено – име на лице (*Мария*).

3.1.2. Съставно съществително собствено – име на лице (*Мария Петрова*).

3.1.3. Съставно съществително собствено – име на лице, съдържащо категоризиращо нарицателно име (*Панеса Йюана*). Случаи, в които категоризиращото нарицателно се е утвърдило като част от името, са много редки – предимно функционират като прозвища на известни личности.

3.1.4. Словосъчетание с опора съществително собствено име – име на лице, разширено с категоризиращо нарицателно име (*чичо Иван, професор Иванов*).

3.1.5. Словосъчетание с опора съставно съществително собствено – име на лице, разширено с категоризиращо нарицателно име (*професор Йордан Пенчев*).

3.2. Имена на географски обекти

3.2.1. Съществително собствено – географско название (*Дунав, Вашингтон*). Съществителното собствено – географско название, може да съвпада с лично име (*парк „Линкълн“*), много по-рядко – с име на организация (*площад „Херълд“*).

3.2.2. Съставно съществително собствено – географско название (*Черна гора, Сан Антонио*). В състава му може да влиза и лично име (*Ел Салвадор*), много по-рядко – име на организация.

3.2.3. Съставно съществително собствено – географско название, съдържащо категоризиращо нарицателно име (*Стара планина, Мандренско езеро*). Категоризиращото нарицателно име не може да бъде изпуснато или повто-

рено в словосъчетание с името. Съставното име може да включва лично име (*Форт Кристина*).

3.2.4. Словосъчетание с опора съществително собствено име – географско название, разширено с категоризиращо нарицателно име (*река Дунав*).

3.2.5. Словосъчетание с опора съставно съществително собствено – географско название, разширено с категоризиращо нарицателно име (*вулкан Охос дел Саладо*).

3.3. Названия на организации

3.3.1. Съществително собствено – название на организация (*„Гугъл“*). Може да съвпада с лично име (*„Касио“*) или с географско название (*„Ню-киа“*).

3.3.2. Съставно съществително собствено – название на организация (*„Пежо Ситроен“*). Може да съвпада с или да съдържа лично име (*„Джон Диъри“*) или географско название (*„Ню Йорк таймс“*).

3.3.3. Съставно съществително собствено – название на организация, съдържащо категоризиращо нарицателно име (*Европейска банка за възстановяване и развитие*). Категоризиращото нарицателно име не може да бъде повторено или изпуснато. Съставното название на организация може да съдържа лично име или географско название (*Университетът в Кеймбридж*).

3.3.4. Словосъчетание с опора съществително собствено име – название на организация, разширено с категоризиращо нарицателно име (*корпорация „Развитие“ КДА*).

3.3.5. Словосъчетание с опора съставно съществително собствено – название на организация, разширено с категоризиращо нарицателно име (*компанията „Бизнес модели“ АД*).

Семантичната класификация е мотивирана от семантичната структура (и кореспондиращата ѝ синтактична структура) на подчинителните словосъчетания, които се образуват с опора собствено име. Семантичната структура за словосъчетанията с опора лично име се определя от характеристиките на общия референт на съществителното собствено име – лице, и характеристиките на референта, дефиниран от категоризиращото нарицателно име. Семантичната структура за географските названия и названията на организации се определя само от характеристиките на референта, определен от категоризиращото нарицателно име (което може да бъде както експлицитно, така и имплицитно изразено). За целите на по-подробното описание и категоризиране на членовете на семантичната класификация формулирахме следните семантични подкласове за трите основни семантични класа (лица, организации, географски обекти):

3.4. Имена на длъжности, географски обекти и организации

3.4.1. Имена на лица

Подкласове: длъжности в държавното управление (*министър, министър-председател*); изпълнителни длъжности в компании (*директор, изпълнителен директор*); длъжности в съдебната власт (*съдия, съдия-следовател*); академични длъжности (*доцент, главен асистент*); военни чинове (*полковник, генерал-майор*); професия (*инженер*); академични титли (*доктор*); официални обръщения (*господин*); аристократични титли (*княгиня*); религиозни чинове (*епископ*); роднински названия (*сестра*); титли на светии (*свети*).

3.4.2. Имена на географски обекти:

Подкласове: природни обекти (*река, планинска верига*) и административни обекти (*град, булевард*).

3.4.3. Имена на организации:

Подкласове: за имена на организации: църковни организации (*църква, абатство*); правителствени организации (*министерство, държавна агенция*), бизнес организации (*компания, събирателно дружество*); политически организации (*партия, движение*); обществени комерсиални обекти (*ресторант, банка*); обществени граждански обекти (*музей, етнографска къща*); медийни организации (*печатница, телевизия*); неформална граждански организации (*асоциация, неправителствена организация*); изпълнителски групи и групи от творческия сектор (*оркестър, хорова капела*).

Тъй като един човек може да заема различни длъжности и да изпълнява различни функции, личното име може да бъде определено с повече от едно категоризиращо нарицателно име, принадлежащо към различни подкласове. Някои категоризиращи нарицателни имена не могат да се комбинират в определен културно-социален контекст – например едновременно заемане на длъжности в изпълнителната и в съдебната власт в определени системи на държавно управление (ограниченията върху комбинаториката на наименованията и категоризиращите нарицателни имена обаче не са универсални и подлежат на промени в зависимост от културния контекст).

В рамките на нашето изследване групирахме семантичните подкласове като: (I) позиции / длъжности в (социална) йерархия – те не са свързани с определено лице, но лицето може временно да ги заема (*министър, главен изпълнителен директор*); (II) титли за квалификации (*доктор, инж., арх.*); (III) почетни звания (*господин, сър, Ваше Светейшество*); (IV) роднински названия (*братовчед*). Категоризиращите нарицателни имена обикновено се подреждат в определен ред (спрямо опората), като тези от (II) стоят по-близо до личното име в сравнение с тези в (I): *министърът арх. Иванов*. Роднинските названия (от група IV) се съчетават с почти всички други групи,

но са разположени най-вляво в именната фраза и най-често са придружени от конституент за изразяване на принадлежност: *брат ми, професор доктор Попов*. Почетните звания може да се срещнат преди названията за длъжности от държавното управление, изпълнителните длъжности в бизнеса, религиозните чинове и академичните длъжности и титли, но са употребени самостоятелно при обръщение, като в тези случаи са в позиция, която е външна за същинската структура на именната фраза (*Госпожо министър (Павлова), Господин премиер; но министър-председателят господин Борисов*). Названията за длъжности в държавното управление, изпълнителните длъжности в бизнеса и религиозните чинове обикновено не се срещат заедно, като обаче могат да предшестват академични длъжности и титли. Академичните позиции винаги предшестват академичните титли. Съкращенията изискват фраза със собствено име в опората и се появяват колкото е възможно по-близо до нея. В 1. е дадено схематично представяне на възможните комбинации 1.а., заедно с примери 1.б.

1.а. (Държавно управление | Изпълнителни длъжности | Религиозни чинове) > Военни чинове > [Академични длъжности | Академични титли | (Академични длъжности > Академични титли)] > Квалификация/Професия

1.б. Примери:

премиерът генерал-майор проф. Иванов
заместник-министър на отбраната генерал-лейтенант инж. Атанас Запрянов
министърът на земеделието и храните госпожа Десислава Танева
архиепископът на Амалфи и Кава (Де Тирени) господин Орацио Соричели
изпълнителният директор на „Миньор“ генерал Йордан Димитров

4. Езикови ресурси за автоматично разпознаване на именувани обекти

Едно от приложенията на класификацията на именуваните обекти и описанието на специфичните семантични и граматични характеристики на различните класове именувани обекти и структурите, които те изграждат, е свързано със създаването на ресурси за автоматичното им разпознаване. Методите за автоматично разпознаване на именувани обекти се разделят най-общо на два вида. Първата група методи, които са и по-широко използвани в по-ранното развитие на тази област, използват правила за описание и разпознаване на именуваните обекти. Правилата са лингвистични (именуваните обекти са езикови обекти) и с тях се описват: графичното представяне на наименованията (буквени и числови символи, пунктуация) и характеристиките на елементите в състава на наименованието. Втората група методи се основават на статистически (количествени) модели, като се използват

алгоритми от различен вид – управлявано обучение (supervised learning), частично управлявано обучение (semi-supervised learning), неуправлявано обучение (unsupervised learning)¹, както и хибридни методи и комбинации от различни алгоритми в рамките на една система за автоматично разпознаване. Най-често използваните алгоритми в категорията управлявано обучение се „обучават“ да разпознават и класифицират именуваните обекти (или които и да било други обекти) върху вече аотирани езикови данни – корпуси, в които именуваните обекти са класифицирани от лингвисти или аотацията е приписана автоматично, например с помощта на ресурси като Уикипедия. За български Стоянова (Стоянова/Stoyanova 2012, 2014) предлага метод за автоматична категоризация на различни видове лексикални съставни единици, включително съставни именувани обекти, като използва класификация на съставните лексикални единици въз основа на тяхната идиоматичност (по Болдуин и др./Baldwin et al. 2003) и информация за състава и графичните характеристики на конституентите на съставната лексикална единица. Методът е тестван върху корпус от статии от Уикипедия и използва голям речник на съставните лексикални единици, компилиран автоматично и проверен.

Метод за разпознаване на именувани обекти в български, основан на правила, е предложен от Осенова и Колковска (Osanova, Kolkovska 2002). Правилата, част от които са и речници, са внедрени в модула за разпознаване на именувани обекти БулНейм (BulName), част от системата БулТриБанк, който включва граматика, основана на правила, и приложение за разпознаване на частични съвпадения. Методът е насочен предимно към разпознаване на имената, като се използват две граматика за именната фраза – граматика за разпознаване на именувани обекти и граматика за разпознаване на поредица от думи (чънкиране), които са част от БулТриБанк. Разработката обаче не предлага по-подробно описание на семантичната и синтактичната структура на именуваните обекти и зависимостите между двете.

Една от задачите в рамките на настоящото изследване включва създаване на езикови ресурси за приложение при методи за автоматично разпознаване на именувани обекти, основани на предварително формулирани правила. Подолу накратко ще опишем създадените езикови ресурси – речници и малък корпус с аотирани именувани обекти от трите класа и съответните им подкласове.

¹ Алтернативни термини в български за „управлявано обучение“ и „неуправлявано обучение“ са съответно „обучение с учител“ и „обучение без учител“.

4.1. Речници

В рамките на изследването бяха съставени следните речници:

4.1.1. Речници с имена на лица

Бяха съставени речници със собствени лични имена за лица от мъжки пол (5536 единици) и от женски пол (6483 единици) и съответните фамилни имена (за лица от мъжки пол – 3455 единици, и за лица от женски пол – 3419 единици), включително речници с обръщения към лица от мъжки пол, както и речници със съществителните нарицателни, които принадлежат към някоя от формулираните категории, с определени и неопределени форми, за лица от женски и мъжки пол, както следва (посочена е само бройката на нечленуваните форми (създадени са и речници за членуваните форми) и не е посочена бройката на възможните съкращения):

А) Академични длъжности, напр. *доцент, главен асистент, проф.* (едносъставни, м.р.: 9; едносъставни, ж.р.: 8; съставни, м.р.: 15; съставни, ж.р.: 15).

Б) Академични титли, напр. *доктор, д-р* (едносъставни, м.р.: 2; едносъставни, ж.р.: 1; съставни, м.р.: 70; съставни, ж.р.: 23).

В) Длъжности в държавното управление, напр. *министър, генерален консул, извънреден и пълномощен посланик* (едносъставни, м.р.: 94; едносъставни, ж.р.: 48; съставни, м.р.: 115; съставни, ж.р.: 105).

Г) Изпълнителни длъжности в бизнес организация, напр. *мениджър, главен изпълнителен директор* (едносъставни, м.р.: 33; едносъставни, ж.р.: 33; съставни, м.р.: 146; съставни, ж.р.: 149).

Д) Военни чинове, напр. *генерал, генерал от запаса, запасен генерал* (едносъставни, м.р.: 40; едносъставни, ж.р.: 9; съставни, м.р.: 158).

Е) Аристократични титли, напр. *виконт, графиня* (едносъставни, м.р.: 66; едносъставни, ж.р.: 33; съставни, м.р.: 4; съставни, ж.р.: 9).

Ж) Религиозни чинове, напр. *архиепископ, игуменка* (едносъставни, м.р.: 94; едносъставни, ж.р.: 7; съставни, м.р.: 15; съставни, ж.р.: 1).

З) Длъжности в съдебната система, напр. *съдия, главен прокурор* (едносъставни, м.р.: 31; едносъставни, ж.р.: 29; съставни, м.р.: 110; съставни, ж.р.: 113).

И) Професии, напр. *готвач, печатар мантограф* (според Националната класификация на професиите и длъжностите) (едносъставни, м.р.: 1386; едносъставни, ж.р.: 1355; съставни, м.р.: 1642; съставни, ж.р.: 1734).

Й) Роднински названия, напр. *сестра, първи братовчед* (едносъставни, м.р.: 27; едносъставни, ж.р.: 26; съставни, м.р.: 10; съставни, ж.р.: 9).

4.1.2. Речници със съществителни нарицателни, категоризиращи географски обекти (362 единици).

А) Природни обекти, напр. *карстово езеро, екозона* (едносъставни, м.р.: 72; едносъставни, ж.р.: 38; едносъставни, ср.р.: 17; съставни, м.р.: 14; съставни, ж.р.: 16; съставни, ср.р.: 5).

Б) Административни обекти, напр. *столица*, вкл. инфраструктурни, напр. *железопътна гара* (едносъставни, м.р.: 21; едносъставни, ж.р.: 22; едносъставни, ср.р.: 17; съставни, м.р.: 49; съставни, ж.р.: 71; съставни, ср.р.: 20).

4.1.3. Речници със съществителни нарицателни, категоризиращи имена на организации (819), както следва:

А) Църковни организации, напр. *митрополия* (едносъставни, м.р.: 3; едносъставни, ж.р.: 14; едносъставни, ср.р.: 5; съставни, м.р.: 15; съставни, ж.р.: 24; съставни, ср.р.: 10).

Б) Правителствени организации, напр. *министерство*, *Министерски съвет* (едносъставни, м.р.: 14; едносъставни, ж.р.: 21; едносъставни, ср.р.: 9; съставни, м.р.: 10; съставни, ж.р.: 16; съставни, ср.р.: 5).

В) Политически организации, напр. *партия*, *сдружение* (едносъставни, м.р.: 11; едносъставни, ж.р.: 17; едносъставни, ср.р.: 5; съставни, м.р.: 2; съставни, ж.р.: 6; съставни, ср.р.: 1).

Г) Търговски организации, напр. *компания*, *модна къща*, *брокерска къща* (едносъставни, м.р.: 35; едносъставни, ж.р.: 18; едносъставни, ср.р.: 7; съставни, м.р.: 14; съставни, ж.р.: 26; съставни, ср.р.: 11).

Д) Търговски обекти, напр. *ресторант*, *месарница* (едносъставни, м.р.: 41; едносъставни, ж.р.: 49; едносъставни, ср.р.: 10; съставни, м.р.: 37; съставни, ж.р.: 2; съставни, ср.р.: 15).

Е) Обществени организации, напр. *училище*, *етнографска къща*, *музеен комплекс* (едносъставни, м.р.: 26; едносъставни, ж.р.: 29; едносъставни, ср.р.: 6; съставни, м.р.: 24; съставни, ж.р.: 17; съставни, ср.р.: 1).

Ж) Издателски организации, напр. *вестник*, *издателска къща* (едносъставни, м.р.: 33; едносъставни, ж.р.: 54; едносъставни, ср.р.: 14; съставни, м.р.: 9; съставни, ж.р.: 10; съставни, ср.р.: 13).

З) Граждански организации, напр. *неправителствена организация*, *сдружение* (едносъставни, м.р.: 2; едносъставни, ж.р.: 4; едносъставни, ср.р.: 4; съставни, м.р.: 10; съставни, ж.р.: 15; съставни, ср.р.: 7).

И) Изпълнителски групи, напр. *балет*, *оперна трупа* (едносъставни, м.р.: 22; едносъставни, ж.р.: 16; едносъставни, ср.р.: 5; съставни, м.р.: 12; съставни, ж.р.: 26; съставни, ср.р.: 7).

4.2. Корпус

В рамките на изследването и за изясняване на описаните структури беше създаден корпус с текстове от Българския национален корпус (художествена литература; публицистика и новини; научна и научнопопулярна от областите химия, история, икономика; субтитри), с обем 101 507 токъна².

² С термина „токън“ се означава поредица от символи между два отделящи символа като шпация и/или пунктуационен знак.

Собствените имена бяха анотирани ръчно, като корпусът беше допълнен с различни именувани обекти, с различна структура и (комбинации от) конституенти (опора, модификатори, клитики), графични съкращения и др. Анотирани са именуваните обекти и основните им конституенти – главна част: лични имена (*Симеон Дянков* в пример 2.а.; *Дунав* в пример 2.б.; *Колеж по застраховане и финанси* и *Берн* в пример (2.в.); приложения, вкл. опорите и семантичните класове – с отделен таг за всеки семантичен клас (напр. в пример (2.а.) <2-1> за длъжности в държавното управление, като <2-101> е за членуваната форма; в 2.б. <21-2> – за географски обект, ж.р., нечленуван; <21> – име на географски обект; а в пример 2.в. <22-1> – за административен/неприроден обект, м.р., нечленуван). Зависимите конституенти не се отбелязват, ако не са част от името (напр. *Колеж по застраховане и финанси* в пример 2.в., но цялата фраза се анотира в общ таг <2m> – в пример 2.а.; при координирани фрази координираните елементи се анотират в общ таг <2c> в пример 2.а.

2.а. <per><2c><2-101>вицепремиера</2-1>и<2m><2-1>министър</2-1> на финансите</2m></2c> <1-1>Симеон</1-1><1-9>Дянков</1-9></per>

2.б. <geo><21-2>река</21-2><21>Дунав</21></geo>

2.в. <org><35><35-1>Колеж</35-1> по застраховане и финанси</35> в <geo><22-1>град</22-1> <22>Берн</22></geo></org>

В момента корпусът съдържа: именувани обекти с референт лице: 721 (лично име: 319; фамилно име: 369; професии: 86; роднински названия: 70; титли: 53; длъжности в изпълнителната власт: 33; академични длъжности: 21; академични титли: 18; и др.); именувани обекти с референт географски обект: 151 (природни обекти: 15; административни: 144; имена: едносъставни: 116; съкращения: 8; съставни: 25); именувани обекти с референт организация: 184 (църковни организации: 1; правителствени организации: 69; политически организации: 13; търговски организации: 23; търговски обекти: 4; обществени организации: 38; издателски организации: 4; граждански организации: 29; изпълнителски групи: 2).

Подобни лингвистични ресурси могат да се използват за трениране и формулиране и приложение на правила за разпознаване на именуваните обекти в текст.

В следващите раздели ще представим описание на различните класове и подкласове чрез: езиково независими семантични структури (модели) за описание на характеристиките на имената на лица, имената на географски обекти и имената на организации; и езиково специфични синтактични структури – именни фрази.

5. Семантичен модел за описание на именуваните обекти³

Предлагаме модел за описание на личните имена, географските имена и имената на организации посредством: 1) формулиране на езиково независими семантични структури за описание; 2) представяне на езиково специфични синтактични структури, посредством които се реализират семантичните структури.

Семантичните структури се разграничават въз основа на: а) класа на собственото име; б) класа на категоризиращата дума; в) вида на категоризиращата дума: вътрешна или външна по отношение на собственото име; г) семантичната структура, която се асоциира с класа собствени имена.

5.1. Класове собствени имена

Идеята, която стои в основата на разделянето на собствените имена на класове, е, че различни собствени имена се комбинират с различни категоризиращи думи, което се определя от семантиката на референта на собственото име, например личните имена се използват за назоваване на близки хора (*баба Мария*); фамилните имена се използват за назоваване на хора в ролята им на заемащи определена административна, професионална или обществена функция (*инженер Коев*, но не **дядо Коев*) и т.н. Класовете собствени имена, които назовават различни класове референти и съответно се комбинират с различни семантични класове категоризиращи думи, са, както следва: лично име; лично име, разширено с идентификатор (*Инокентий V*; *Боне Крайне-неца*); лично име, разширено с фамилно име (или комбинация от бащино и фамилно име); фамилно име; име на географски обект; име на географски обект – съставна дума с вътрешна категоризираща дума (*Индийски океан*); име на организация; име на организация – съставна дума с вътрешна категоризираща дума (*Асоциация на българските застрахователи*).

5.2. Категоризиращи думи

За да се интерпретира правилно дадено собствено име, трябва да е известен неговият референт (или експлицитно, или от контекста). Референтът е конкретен обект и може да се характеризира от нарицателно име (*професор, река, театър*), което назовава класа (категоризиращия обект), към който принадлежи конкретният обект. Характеристиките на референта са прояви на характеристиките на класа, които са валидни за референта. Класът, към който принадлежи конкретният обект, назован със собствено име, може да се свързва с различни собствени имена, за да се назовават различни конкретни

³ Автор на пета, шеста и седма част от студията е Светла Коева.

обекти: *река Марица, река Янтра* и т.н. Така, както изразът *масата в хола на Иванови* назовава единичен обект, наследяващ конкретни прояви на характеристиките на класа *маса*, така и изразът *Иван, синът на семейство Иванови*, назовава единичен обект, наследяващ прояви на характеристиките на класа *син* и неговите надредни класове: *мъж, човек*.

На семантично равнище съществителното собствено може да се характеризира с признаци, отнасящи се до: а) най-общия клас, към който принадлежи референтът (съответно *човек, място, организация*); б) подредни на общия клас класове, към които принадлежи референтът (например *жена, воден басейн, медия*); в) конкретния клас, с който се характеризира референтът (например *лекарка, река, радио програма*). Отделните признаци характеризират референта по специфичен начин (например *стар, минерален, новосформиран*), като част от тях са семантично задължителни, за да се изрази значението на категоризиращата дума (*професор по физика, доктор на биологическите науки*).

Категоризиращата дума специфицира семантичния клас, към който принадлежи референтът на собственото име. При личните имена едно и също лице може да изпълнява различни функции и да бъде характеризирано с различни категоризиращи думи (*писателят Иван Вазов, възрожденецът Иван Вазов, министърът на образованието Иван Вазов*). Категоризиращите думи могат да се комбинират, ако комбинацията отразява реални отношения (*известният писател и общественик Иван Вазов*). При географските названия и организациите за разлика от личните имена собственото име се характеризира с точно определена категоризираща дума и рядко може да се използват нейните синоними или хипероними (*пещерата Божиите очи, природното образуване Божиите очи, Застрахователно акционерно дружество „Армеец“, компанията „Армеец“*).

Ако категоризиращото име не е експлицитно назовано, то винаги се предполага от контекста, в който се употребява собственото име, в противен случай изказването ще е многозначно. Подобно разбиране за собствените имена е най-близко до теорията за пресупозициите (Камп, Рейл/Kamp, Reyle 1993), която разглежда собствените имена (подобно на местоименията) с функция да препращат към референти, които вече са известни в дискурса.

Следователно, тъй като всяко собствено име реферира към конкретен обект, обектите с общи семантични свойства могат да бъдат групирани в семантични структури, които описват тези свойства. Например фамилните имена се съчетават с категоризиращи нарицателни имена за длъжности в държавната власт, изпълнителни длъжности в компании, длъжности в съдебната система, академични длъжности, академични титли, военни чинове, религиозни титли и професии. Обособени са следните групи категоризиращи думи в зависимост от съчетаемостта им с класове собствени имена и прилежащата им семантична структура: светски длъжности, професии и военни чинове; академични титли; аристократични титли и религиозни длъжности; титли на светци; роднински названия; обръщения; географски обекти; организации.

5.3. Външни и вътрешни категоризиращи думи

При географските названия и организациите категоризиращата дума може да стане част от собственото име (*Черно море*, *Стара планина*, *Търговска банка*) – означаваме я с термина **вътрешна категоризираща дума**. Рядко при собствените имена категоризиращата дума се използва винаги заедно със собственото име (*Принц Чарлз*).

В съставните имена, които съдържат вътрешна категоризираща дума (*Индийски океан*), категоризиращата дума е винаги експлицитна (Карол/Саргол 1985, 144). Ако името съдържа вътрешна категоризираща дума, обикновено не се употребява външна категоризираща дума, която е повторение (**океанът Атлантически океан*, но *село Ново село*), и рядко се употребява различна външна категоризираща дума (*водният басейн Средиземно море*). Съставните имена обаче могат да съдържат и категоризираща дума, която е част от името, но не назовава категоризиращия обект (*читалище „Стара планина“*), в такъв случай употребата на външното категоризиращо име е до голяма степен задължителна, за да се идентифицира еднозначно обектът, назован със собственото име.

В случаите, когато към съставни имена с вътрешна категоризираща дума има и външна категоризираща дума, семантичната и синтактичната структура, които се асоциират с тях, са същите, както при простите имена с външна категоризираща дума ([*новото*] *читалище* [*в квартал „Младост“*] *„Стара планина“*; [*красивата*] *планина* [*в Западна България*] *Люлин*).

Ако съставните имена с вътрешна категоризираща дума нямат външна категоризираща дума, вътрешната категоризираща дума определя задължителните и незадължителните модификатори, които се асоциират с името ([*новата*] *„Пощенска банка“* [*в Атина*])). В такъв случай предпоставените и задпоставените модификатори на категоризиращата дума стават предпоставени и задпоставени модификатори на собственото име.

Вътрешната категоризираща дума може да реализира в собственото име някои от задължителните или незадължителните си модификатори (*организацията* [*Българска*] *асоциация* [*за компютърна лингвистика*])). Ако модификатори на вътрешната категоризираща дума са станали част от името, модификатори от същия вид не могат да се повторят, ако няма различна външна категоризираща дума (**Българска асоциация за компютърна лингвистика по физика*; *Българска асоциация за компютърна лингвистика към Българската академия на науките*; *Сдружението за хора в неравностойно положение* *Българска асоциация за лица с интелектуални затруднения*).

Наличието на външни и вътрешни категоризиращи думи предопределя семантичната структура по следния начин: вътрешната категоризираща дума може да бъде разширена с някои от модификаторите си като част от името и определя модификаторите на собственото име, ако няма външна категоризираща дума.

5.4. Семантични модели (структури)

В семантичните модели са представени възможните комбинации между имената (лица, географски обекти, организации) с различна структура (прос-ти думи, съставни думи, съставни думи с вътрешна категоризираща дума) и семантичните класове на категоризиращите нарицателни имена. Семантичните модели показват също семантичната съчетаемост на категоризиращото нарицателно име с разширения от различни семантични класове. Например лично собствено име може да бъде разширено с роднинско название (*баба Мария*), а роднинското название от своя страна може да бъде специфицирано по различни начини, например по отношение на притежател (*бабата на Даниел Мария*).

Тъй като категоризиращите думи назовават понятия, семантичните отношения, в които те участват, са универсални и важат за еквивалентно понятие, назовано в произволен език. Следователно семантичните модели представят езиково неутрални отношения и могат да се приемат като универсални структури, към които се отнасят специфични синтактични структури за отделните езици.

В литературата се използва различна терминология, за да се характеризират зависимите части в подчинително словосъчетание с главна част съществително: модификатори, квантификатори, спецификатори, детерминатори, компленти, адюнкти, при това някои от тези термини се използват с различно значение. По тази причина ще опишем накратко значението, което влагаме в използваната терминология, за нуждите на това изследване.

С термина **модификатори** ще означаваме всички зависими части в рамките на словосъчетание с главна част съществително. Модификаторите разделяме на детерминатори, квантификатори, спецификатори, компленти и адюнкти. **Квантификаторите** са думи, означаващи (най-общо казано) количествена характеристика по отношение на главното съществително (*първи, втори, всеки* и др.). Те не могат да се съчетават със собствени имена (освен в редки случаи, когато два или повече обекта от един клас са назовани с едно и също име: *втората компания „Тетраком“*, *всяко едно от двете читалища „Светлина“*). **Детерминаторите** са определителният член, показателните и притежателните местоимения. Известно е, че определителният член и показателното местоимение са в дистрибутивна употреба, докато притежателното местоимение се съчетава и с определителния член, и с показателното местоимение. Трябва да се отбележи, че макар и възможна, употребата на показателни местоимения в словосъчетания с главна част собствено име е рядка. Това е следствие от рефериращата функция на собствените имена, която интерпретираме като аналогична на рефериращата функция на личните местоимения. (Изречението *Харесвам тази баба Мария* означава, че има поне две баби, наречени Мария, или че говорещият изразява специално отношение.) Приема се, че квантификаторите и детерминаторите определят референтния статус на

думата или словосъчетанието. **Комплементите** са думи или словосъчетания, лицензирани от опората, което означава, че са необходими, за да се изрази значението на опората (Пейн/Пауе 2011). Следователно комплементите са семантично задължителни, макар в някои употреби да остават синтактично имплицитни, ако се подразбират от непосредствения или широкия контекст. Например словосъчетанията *професорът [по физика] Иванов* и *професор Иванов* съдържат комплемент към категоризиращата дума *професор* (ако се говори за един и същи професор), макар че във втория случай комплементът е имплицитен. **Спецификаторите** са предпоставени разширения – прилагателни или съществителни имена, които конкретизират значението на опорното съществително, без да са семантично и синтактично задължителни ([[новият професор] Иванов]). Множеството от допустими спецификатори се определя от семантиката на референта на съществителното име (професорът може да е млад или стар, скалата – остра или ронлива, Иванов – професор или музикант). **Адюнктите** са задпоставени разширения – предложни групи, които модифицират значението на опората, без да са семантично и синтактично задължителни (*професорът [от София] Йордан Пенчев*). Множеството от допустими адюнкти се определя от семантиката на референта на съществителното име и характеристиките на контекста, в който се реализира. Някои адюнкти – предложни групи, се редуват с предпоставени прилагателни имена (*професорът от Дъблин – дъблинският професор*). Както собственото име, така и категоризиращото име могат да имат собствени спецификатори и адюнкти, а категоризиращото име – и комплементи.

За дефинирането на семантичните модели е необходимо да се намери процедура за разграничаване между комплементи и адюнкти при категоризиращите нарицателни имена, тъй като и комплементите, и адюнктите се изразяват с предложни групи. Прилагат се различни разграничителни тестове между комплементи и адюнкти, повечето от които при глаголи (Радфорд/Radford 1988; Куликовер, Джакендоф/Culicover, Jackendoff 2005; Пейн/Пауе 2010; Хуанг/Hwang 2011 и др.).

По-долу са разгледани някои от процедурите за разграничаване, приложени към опора – съществително име. Например използва се зависимостта, че комплементите се разполагат по-близо до опората от адюнктите. В следния пример предложната фраза *по физика* е комплемент, тъй като отдалечаването му от опората прави израза неграматичен.

3.а. *Професорът [по физика] [със странен външен вид] Иванов.*

3.б. **Професорът [със странен външен вид] [по физика] Иванов.*

Словоредът може да се използва, за да се определи редът на конституентите с йерархия на семантичната връзка и зависимостта на появата на един конституент от друг. Следващите примери показват, че комплементът, изразяващ *област на приложение*, предхожда комплементата, показващ *принадлежност към институция*, а комплементът за *притежание* предхожда комплементата за *принадлежност към институция*.

4.а. *Професорът [по физика] [от Университета в Лондон] Джон Смит.*

4.б. **Професорът [от Университета в Лондон] [по физика] Джон Смит.*

4.в. *Професорът [от Университета в Лондон].*

5.а. *Преподавателят [на моя син] [в Бирмингамския университет] Джон Синклер.*

5.б. **Преподавателят [в Бирмингамския университет] [на моя син] Джон Синклер.*

При категоризиращите нарицателни имена – роднински названия, за да има конституент, означаващ *произход*, трябва да има експлицитно изразен конституент за *притежател* или *притежателно местоимение* (*дъщерята на Иван от Пловдив Мария, *дъщерята от Пловдив Мария*).

Друго правило за комплементност е, че адюнктите от един и същи вид могат да се повтарят, а комплементите – не (както е показано в следващия пример). Това правило е преформулировка на правилото, че адюнктите от един и същи вид могат да образуват съчинителни словосъчетания, в редки случаи – комплементите също, но комплемент и адюнкт не могат да образуват съчинително словосъчетание. Правилото за съчинителното свързване показва, че две фрази са комплементи или адюнкти, но не е достатъчно, за да се определи видът на всяка една от фразите.

6.а. *Професорът [със странен вид] и [странно поведение] Пенчев.* (адюнкти)

6.б. *Професорът [по стилистика и синтаксис] Пенчев.* (комплементи)

6.в. **Професорът [със странен вид] и [по синтаксис] Пенчев.* (адюнкт и комплемент)

Най-подходящо за изследователската ни цел е правилото, с което се проверява дали комплементът и главното съществително едновременно описват един обект, а адюнктът и главното съществително – два различни обекта. За проверка комплементите и адюнктите на категоризиращата дума се включват в сложно съчинено изречение, двете прости изречения на което са копулативни със субект собственото име и с предикат в първото изречение – категоризиращото нарицателно име, а във второто – проверяваната фраза.

7.а. *Професорът [по химия] [със странно поведение] Иван Иванов.*

7.б. **Иван Иванов е [професор] и (той) е [по химия].* (комплемент на категоризиращото име)

7.в. *Иван Иванов е [професор] и (той) е [професор по химия].*

7.г. *Иван Иванов е [професор] и (той) е [със странно поведение].* (адюнкт на категоризиращото име)

Примерът 7.б. е неграматичен, тъй като комплементът и неговото опорно съществително са разделени в две различни прости изречения, за разлика от 7.в. Следователно *по химия* е комплемент към *професор*. За адюнктите изискването опората, към която се отнасят, да е в същото просто изречение, не е задължително, тъй като опората и адюнктите описват различни обекти от действителността.

Потенциалните комплементи и адюнкти, които ще изследваме, са със следната семантика: област на приложение (*асистент* [по право]); принадлежност (към институция или обект) (*асистент* [в Пловдивския университет], (*проходът* „Вратцата“ [във Врачанския балкан]), притежател (*асистентът* [на студентите], *лялята* [на Мая]), местоположение (*асистентът по право* [в Пловдив], *вестник* „Марица“ [в Пловдив], *река Марица* [край Пловдив]), произход (*асистентът* [от Лондон]), спецификация (*студентът* [с тъмна коса]), като, както ще видим от примерите по-долу, *произход* и *спецификация* са винаги адюнкти.

В семантичните модели описваме семантичния обхват на спецификаторите, комплементите и адюнкти, а в синтактичните – синтактичната реализация на детерминаторите, спецификаторите, комплементите и адюнкти. Следва семантично описание на структурата на лични имена, географски имена и имена на организации в семантични модели, които отразяват техните споделени свойства. Информацията в семантичните модели е в реда: клас на съществителното собствено, семантичен клас на категоризиращото нарицателно име, вид на категоризиращото име – външно или вътрешно, семантичен модел и примери.

5.5. Представяне на семантичните модели

5.5.1. Семантичен модел I

(а) Клас на съществителното собствено име: фамилно име; лично име, разширено с фамилно име (ако има фамилно име, може да има и бащино).

(б) Семантичен клас на категоризиращото нарицателно име: изпълнителни длъжности в организации, академични длъжности, длъжности в съдебната власт, длъжности в държавното управление, военни чинове, професии.

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто, сложно (например *министър-председател*) и съставно съществително (например *главен асистент*).

(г) Семантичен модел:

$S^4 = (\text{спецификатор})^{*5} - \text{категоризиращо нарицателно име}^6 - K = (\text{област на приложение})? - K = (\text{притежател})? - (K = (\text{принадлежност към институция}) \text{ или } K = (\text{местоположение}))? - KC = (\text{произход})? - A = (\text{спецификация})^* - \text{собствено име}$

(д) Примери, които илюстрират семантичния модел:

⁴ С означава спецификатор, К – комплемент, А – адюнкт, КС – комплемент на собственото име.

⁵ * означава, че предходният елемент се среща нула, един или повече от един път, ? – нула или един път.

⁶ Категоризиращото нарицателно име е винаги спецификатор на съществителното собствено.

8.а. *Новият заместник-ректор* [К=(област на приложение) по учебната част] [К=(принадлежност) към Аграрния университет в Пловдив] Иван Иванов.

8.б. **Иван Иванов е заместник-ректор и (той) е по учебната част.* Следователно [по учебната част] е комплемент на категоризиращото име.

8.в. **Иван Иванов е заместник-ректор и (той) е към Аграрния университет в Пловдив.* Следователно [към Аграрния университет в Пловдив] е комплемент на категоризиращото име.

9.а. *Известният преподавател* [К=(притежател) на моя син] [КС=(произход) от Дъблин] Джон Смит.

9.б. **Джон Смит е преподавател и (той) е на моя син.* Следователно предложната група [на моя син] е комплемент на категоризиращото име.

9.в. *Джон Смит е преподавател и (той) е от Дъблин.* Тук е необходим допълнителен тест (**Джон Смит е преподавател и (той) е от Дъблин и Лондон.*), който показва, че [от Дъблин] е комплемент на собственото име.

10.а. *Младият асистент* [К=(област на приложение) по икономика] [К=(местоположение) в Лондон] [А=(спецификация) с неизчерпаема енергия] Иван Иванов.

10.б. *Иван Иванов е асистент и (той) е в Лондон.* Значението 'асистент, който преподава в Лондон' не се предава от изречението в 10.б., значението е 'асистент, който се намира в Лондон с някаква цел'.

10.в. **Иван Иванов е асистент в Лондон и Париж.* Следователно [в Лондон] е комплемент на категоризиращото име. Комплементът *местоположение* може да се появи само като алтернатива на комплемента *принадлежност* или можем да разглеждаме появата му като резултат на елиптична употреба на главната част на комплемента *принадлежност* и заместваща функция на подчинената част на комплемента за местоположение (към *Университета в Лондон*)

10.г. *Иван Иванов е асистент и (той) е с неизчерпаема енергия.* Следователно [с неизчерпаема енергия] е адюнкт на категоризиращото име.

Някои категоризиращи съществителни от тази група имат ограничения за част от комплементите, които приемат. Например при най-високите длъжности в йерархията на академичните организации, съдебната власт, армията и държавното управление комплементи, изразяващи *област на компетентност* и *принадлежност*, не се срещат: **министър-председател на финансите / министър на финансите; *ректор по учебната част / заместник-ректор по учебната част; *главният прокурор на двамата престъпници / съдията на двамата престъпници.* Това показва, че след като се обособят семантичните модели, за всяко категоризиращо име трябва да се посочи към кой модел принадлежи.

При този семантичен модел се забелязват следните зависимости в словоред: ако са употребени едновременно задпоставени комплементи със значение 1) *област на приложение* и 2) *принадлежност към институцията или местоположение*, то редът винаги е изброеният. Редът на задпоставените моди-

фикатори е винаги първо комплементите, после адюнктите. Ако има повече от един задпоставен адюнкт, редът им е свободен. Комплементът за *произход* може да се разполага и след съществителното собствено име (*известният професор Джон Синклер от Бирмингам*), което показва, че се отнася директно към него, а не към категоризиращото име. В такъв случай, ако категоризиращото име няма собствени модификатори, то остава нечленувано (*професор Джон Синклер от Бирмингам*).

5.5.2. Семантичен модел II

(а) Клас на съществителното собствено име: фамилно име; лично име, разширено с фамилно име (ако има фамилно име, може да има и бащино).

(б) Семантичен клас на категоризиращото нарицателно име: академични титли.

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто или съставно съществително име (например *доктор на науките*).

(г) Семантичен модел:

C=(спецификатор)* – категоризиращо нарицателно име – K=(област на приложение)? – K=(принадлежност към институция)? – KC=(произход)? – A=(спецификация)* – собствено име

(д) Примери, които илюстрират семантичния модел:

11.а. [C=(спецификатор) *Разсеяният*] *доктор* [K=(област на приложение) *по философия*] [K=(принадлежност) към *Аграрния университет в Пловдив*] *Иван Иванов*.

11.б. **Иван Иванов е доктор и (той) е по философия*. Следователно предложната група [*по философия*] е комплемент на категоризиращото име. Общата семантика на комплеента е *област на приложение*. Академичните титли *доктор на науките* и *кандидат на науките* са съставни съществителни имена, при които определящото прилагателно към *науките* може да варира в дадено множество: *химическите*, *физическите* и под.

11.в. **Иван Иванов е доктор и (той) е към Аграрния университет в Пловдив*. Следователно предложната група [*към Аграрния университет в Пловдив*] е комплемент на категоризиращото име. Нейната обща семантика е *принадлежност към институция*.

12.а. [C=(спецификатор) *Известният*] *доктор* [K=(област на приложение) *по философия*] [C=(спецификатор) *с неизчерпаема енергия*] [KC=(произход) *от Дъблин*] *Джон Смит*.

12.б. *Джон Смит е доктор и (той) е от Дъблин*.

12.в. **Джон Смит е доктор и (той) е от Дъблин и Лондон*. Следователно предложната група [*от Дъблин*] е комплемент на собственото име.

12.г. *Джон Смит е доктор и (той) е с неизчерпаема енергия*. Следователно предложната група [*с неизчерпаема енергия*] е адюнкт на категоризиращото име.

В този семантичен модел не се наблюдават семантичните отношения *притежател* и *местоположение*. Зависимостите в словоредата са същите, както при семантичен модел I: ако са употребени едновременно задпоставени комплементи, редът е: *област на приложение, принадлежност към институция*. При задпоставените модификатори първо са комплементите, после адюнктите. Ако има повече от един задпоставен адюнкт, редът им е свободен. Комплементът за *произход* може да се разполага и след съществителното собствено име (*известният доктор по философия Джон Смит от Дъблин, доктор Смит от Дъблин*), тъй като пояснява директно него, а не категоризиращото нарицателно име.

5.5.3. Семантичен модел III

(а) Клас на съществителното собствено име: лично име; фамилно име; лично име, разширено с идентификатор; лично име, разширено с фамилно име (рядко).

(б) Семантичен клас на категоризиращото нарицателно име: аристократични титли; религиозни чинове.

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто или съставно съществително (например *Велик херцог*).

(г) Семантичен модел:

$C=(\text{спецификатор})^*$ – категоризиращо нарицателно име – ($K=(\text{принадлежност към институция})$ или $K=(\text{местоположение})$)? – $KC=(\text{произход})$? – $A=(\text{спецификация})^*$ – собствено име

(д) Примери, които илюстрират семантичния модел:

13.а. [$C=(\text{спецификатор})$ *Новият*] *епископ* [$K=(\text{местоположение})$ *в Рим*] *Йосиф Молайони*.

13.б. *Йосиф Молайони е епископ и (той) е в Рим.* (намира се в Рим)

13.в. **Йосиф Молайони е епископ в Рим и Париж.* Следователно предложната група [*в Рим*] е комплемент на категоризиращото име. Нейната обща семантика е *местоположение*.

14.а. [$C=(\text{спецификатор})$ *Новият*] *граф* [$A=(\text{спецификация})$ *с благородно сърце*] *Уилям Бентинк*.

14.б. *Уилям Бентинк е граф и (той) е с благородно сърце.* Следователно предложната група [*с благородно сърце*] е адюнкт на категоризиращото име.

15.а. [$C=(\text{спецификатор})$ *Младият*] *епископ* [$K=(\text{принадлежност към институция})$ *на Българската православна църква*] *Вартоломей*.

15.б. **Вартоломей е епископ и (той) е на Българската православна църква.* Следователно предложната група [*на Българската православна църква*] е комплемент на категоризиращото име. Нейната обща семантика е *принадлежност към институция*.

16.а. [$C=(\text{спецификатор})$ *Младият*] *епископ* [$KC=(\text{произход})$ *от село Чу-чулига*] *Яков*.

16.б. *Яков е епископ и (той) е от село Чучулига. *Яков е епископ и (той) е от село Чучулига и село Ягода.* Следователно предложната група [от село Чучулига] е комплемент на собственото име.

В този семантичен модел не се наблюдават семантичните отношения *притежател* и *област на приложение*. Редът на модификаторите следва модел, при който задпоставеният комплемент (може да има само един комплемент – за *принадлежност към институция* или за *местоположение*) предхожда адюнктите. Адюнктът за *произход* може да се разполага след съществителното собствено име (*Младият епископ Яков от село Чучулига*), тъй като пояснява него, а не категоризиращото нарицателно име.

5.5.4. Семантичен модел IV

(а) Клас на съществителното собствено име: лично име; лично име, разширено с фамилно име (рядко, и то не при всички роднински названия (*братовчед му Иван Тодоров*)).

(б) Семантичен клас на категоризиращото нарицателно име: роднински названия.

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Роднинските названия са прости или (рядко) съставни съществителни имена (например *първи братовчед*).

(г) Семантичен модел:

$S=(\text{спецификатор})^* - \text{категоризиращо нарицателно име} - K=(\text{притежател})? - KS=(\text{произход})? - A=(\text{спецификация})^* - \text{собствено име}$

(д) Примери, които илюстрират семантичния модел:

17.а. [S=(спецификатор) *Чернокосият*] [C=(спецификатор) *пъстроок*] син K=(притежател) *на Мария*] *Даниел*.

17.б. **Даниел е син и (той) е на Мария.* [*На Мария*] е комплемент за *притежател* на категоризиращото име.

18.а. [C=(спецификатор) *Чернокосият*] *му син* [KS=(произход) *от София*] *Даниел*.

18.б. *Даниел е негов син и (той) е от София.*

18.в. **Даниел е негов син и (той) е от София и Пловдив.* [*От София*] е комплемент на собственото име.

19.а. [C=(спецификатор) *Чернокосият*] *му син* [A=(спецификация) *с красиви очи*] *Даниел*.

19.б. *Даниел е негов син и (той) е с красиви очи.* Следователно [*с красиви очи*] е адюнкт на категоризиращото име.

При този семантичен модел не могат да се появят комплементи за *област на приложение* и | или *местоположение*. За да се появи фраза, показваща *произход*, задължително трябва да има експлицитно изразен комплемент за *притежател* или *притежателно местоимение*.

20.а. **Синът от София Даниел.*

20.б. *Синът му от София Даниел.*

20.в. *Синът на Станислав от София Даниел*. (В подобни примери има многозначност – както конституентът [*синът*], така и конституентът [*Станислав*] може да е главна част по отношение на предложната група [*от София*].

Характерно за този семантичен модел е, че има само един комплемент към категоризиращото име и той предхожда адюнктите. Комплементът на собственото име за *произход* се разполага пред или след съществителното собствено име, тъй като е негов модификатор. Появата на комплемента за *произход* зависи от експликацията на комплемента за *притежател* или *притежателно местоимение*.

5.5.5. Семантичен модел V

(а) Клас на съществителното собствено име: лично име; лично име, разширено с фамилно име (рядко); лично име, разширено с идентификатор.

(б) Семантичен клас на категоризиращото нарицателно име: титли на светии (ограничен брой думи).

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто или съставно съществително (например *свети мъченик*).

(г) Семантичен модел:

$C=(\text{спецификатор})^* - \text{категоризиращо нарицателно име} - KC=(\text{произход})?$
 $- A=(\text{спецификация})^* - \text{собствено име}$

(д) Примери, които илюстрират семантичния модел:

21. [$C=(\text{спецификатор})$ *Първият*] [$C=(\text{спецификатор})$ *български*] *свети мъченик Боян-Енравота*.

22.а. *Свети мъченик Илия Чинтуловски* [$KC=(\text{произход})$ *от село Юрта*].

22.б. *Илия Чинтуловски е мъченик и (той) е от село Юрта*.

22.в. **Илия Чинтуловски е мъченик и (той) е (мъченик) от село Юрта и село Ягода*. [*От Юрта*] е комплемент със значение *произход на собственото име*.

22.г. *Свети мъченикът Илия Чинтуловски* [$KC=(\text{произход})$ *от село Юрта*].

Спецификаторите – предпоставени пред категоризиращото име, са редки. Комплементите за *произход* могат да се разполагат пред или след собственото име.

5.5.6. Семантичен модел VI

(а) Клас на съществителното собствено име: фамилно име; лично име, разширено с фамилно име.

(б) Семантичен клас на категоризиращото нарицателно име: официални обръщения (ограничен брой думи).

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто или съставно съществително (например *Негово Превъзходителство*).

(г) Семантичен модел:

C=(спецификатор)* – категоризиращо нарицателно име – собствено име
– КС=(произход)?

(д) Примери, които илюстрират семантичния модел:

23.а. *Господин Иван Иванов.*

23.б. [C=(спецификатор) *Уважаеми*] *господин Иван Иванов.*

24.а. *Госпожа Иванова* [КС=(произход) *от София*].

24.б. **Иванова е госпожа и (тя, госпожата) е от Пловдив и София.* [От *София*] е комплемент на собственото име.

Спецификаторите – предпоставени пред категоризиращото име, са редки. Редки са и задпоставени комплементи за *произход*, които могат да се разполагат само след собственото име, при това, ако фразата не е употребена като обръщение.

5.5.7. Семантичен модел VII

(а) Клас на съществителното собствено име: име на географски обект.

(б) Семантичен клас на категоризиращото нарицателно име: категоризиращо нарицателно име за географски обект (природни обекти, административни обекти).

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто или съставно съществително име (например *природен резерват*).

(г) Семантичен модел:

A=(спецификация)* – категоризиращо нарицателно име – A=(притежател)? – (A=(принадлежност към обект)? или A=(местоположение)?) – A=(спецификация)* – собствено име

(д) Примери, които илюстрират семантичния модел:

25.а. *Малкият град* [A=(принадлежност към обект) *към област Стара Загора*] *Чирпан.*

25.б. *Чирпан е град и (той) е към област Стара Загора.* [Към област *Стара Загора*] е адюнкт.

25.в. *Град Чирпан към област Стара Загора.* Метатезата показва, че адюнктът е към собственото име.

26.а. [C=(спецификатор) *Странджанското*] *село Българи.*

26.б. *Село Българи* [A=(местоположение) *в Странджа*].

26.в. *Българи е село и (то) е в Странджа.* [В *Странджа*] е адюнкт на собственото име.

27.а. *Красивият град* [A=(местоположение) *близо до река Марица*] *Пловдив.*

27.б. *Пловдив е град и (той) е близо до река Марица.* [Близо до река *Марица*] е адюнкт.

27.в. *Красивият град Пловдив* [A=(местоположение) *близо до река Марица*]. Адюнктът е към собственото име.

28.а. *Красивият град* [A=(спецификация) с много тепета] Пловдив.

28.б. *Пловдив е град и (той) е с много тепета.* [С много тепета] е адюнкт на категоризиращото име.

29.а. *Градът* [A=(притежател) на поетите] Стара Загора.

29.б. *Стара Загора е град и (той) е на поетите.* [На поетите] е адюнкт на категоризиращото име.

За разлика от семантичните структури при някои лични имена модификаторите за *притежател*, за *принадлежност* към (природен или административен) *обект* или за *местоположение* са адюнкти. Модификаторът за *притежател* рядко изразява същинско притежание. Адюнктите за *принадлежност* към *обект* (*река Чая от водния басейн на Тракийската низина*) и *местоположение* (*река Чая в Родопите*) могат да се разполагат и след собственото име, което показва, че са негови преки модификатори. Ако категоризиращото име няма модификатори, адюнктите за *принадлежност* към *обект* и *местоположение* не могат да се разполагат пред собственото име (например *хубавата река в Родопите Чая*, но **река в Родопите Чая*).

5.5.8. Семантичен модел VIII

(а) Клас на съществителното собствено име: име на географски обект – съставна дума с вътрешна категоризираща дума.

(б) Семантичен клас на категоризиращото нарицателно име: категоризиращо нарицателно име за географски обект (природни обекти, административни обекти).

(в) Категоризиращото нарицателно име е вътрешно по отношение на собственото име.

(г) Семантичен модел:

C=(спецификатор)* – категоризиращо нарицателно име – A=(притежател)? – (A=(принадлежност към обект)? или A=(местоположение)?) – A=(спецификация)* – собствено име, включващо категоризиращо нарицателно име

(д) Примери, които илюстрират семантичния модел:

30.а. [C=(спецификатор) *Живописният*] *Панчаревски пролом* [A=(местоположение) *на река Искър*].

30.б. *Панчаревският пролом е пролом и (той) е на река Искър.* [На река Искър] е адюнкт към собственото име.

31.а. *Красивата Ягодинска пещера* [A=(спецификация) с *пещерни перли*].

31.б. *Ягодинската пещера е пещера и (тя) е с пещерни перли.* [С *пещерни перли*] е адюнкт на собственото име.

Когато категоризиращото нарицателно име е станало част от собственото име – географско название, и няма употребено друго категоризиращо име, се наблюдават следните възможности:

– Допустими модификатори на категоризиращото нарицателно име се реализират в самото собствено име (*Дяволското гърло*);

– Допустими модификатори на категоризиращото нарицателно име, които не са реализирани в собственото име (*Дяволското гърло*), се разполагат, както следва: спецификаторите преди съставното собствено име, а адюнктите – след него (*триетажната Ягодинска пещера в Смолянска област, красивата Ягодинска пещера с пещерни перли*). Не се наблюдава употреба на адюнкт – притежател.

– Всички реализирани модификатори са към собственото име.

Ако има употребено друго категоризиращо име (*пещерата Орлова чука*), то може да приеме собствени модификатори, отговарящи на семантичния модел, включително *притежател*.

32. *Пещерата* [A=(спецификация) с *подземен водопад*] *Дяволското гърло*. (адюнкт към категоризиращото име)

33. *Пещерата* [A=(притежател) на *прилепите*] *Орлова чука*. (адюнкт към категоризиращото име)

34. *Карстовото понижение в южния склон* [A=(принадлежност към обект) на *Западна Стара планина*] *Раниловско поле*. (адюнкт към собственото име)

35. *Интересното природно образувание* [A=(местоположение) *край река Осъм*] *Деветашката пещера*. (адюнкт към собственото име)

Адюнктите за *принадлежност към обект* (*река Студеното дере от водосборния басейн на Марица*) и *местоположение* (*река Студеното дере в Тракийската низина*) могат да се разполагат и след собственото име, което показва, че са негови преки модификатори. Ако категоризиращото име няма модификатори, адюнктите за *принадлежност към обект* и *местоположение* не могат да се разполагат пред собственото име (например *малката река в Тракийската низина Студеното дере*, но **река в Тракийската низина Студеното дере*).

Семантичните модели VII и VIII са идентични, но моделът, който се отнася до съставни географски имена с вътрешно категоризиращо име (VIII), може да се реализира както по отношение на вътрешното категоризиращо име, така и по отношение на друго, външно, категоризиращо нарицателно име, ако има употребено такова.

5.5.9. Семантичен модел IX

(а) Клас на съществителното собствено име: име на организация.

(б) Семантичен клас на категоризиращото нарицателно име: название на организация (църковна организация, партия, правителствена организация / институция, компания, търговски обект, обществен обект, медийна организация, неправителствена организация, изпълнителска и творческа организация).

(в) Категоризиращото нарицателно име е външно по отношение на собственото име. Може да бъде просто, сложно (например *кандидат-студент*) или съставно съществително име (например *манастирски комплекс*).

(г) Семантичен модел:

C=(спецификатор)* – категоризиращо нарицателно име – K=(област на приложение)? – A=(притежател)? – (A=(принадлежност към обект) – A=(местоположение))?) – A=(спецификация)* – собствено име

(д) Примери, които илюстрират семантичния модел:

36.а. *Новосъздадената фирма* [K=(област на приложение) за производство на чипове] „Сирена“.

36.б. *„Сирена“ е фирма и (тя) е за производство на чипове. [За производство на чипове] е комплемент на категоризиращото име.

37.а. *Новата фирма* [A=(притежател) на Лили Иванова] ЛИ 1 ЕООД.

37.б. ЛИ 1 ЕООД е фирма и (тя) е на Лили Иванова. [На Лили Иванова] е адюнкт.

37.в. *Фирма ЛИ 1 ЕООД на Лили Иванова*. [На Лили Иванова] е адюнкт на собственото име.

38.а. *Новият филиал* [A=(принадлежност към обект) на Пловдивския университет] в Кърджали „Любен Каравелов“.

38.б. „Любен Каравелов“ е филиал и (той) е на Пловдивския университет. [На Пловдивския университет] е адюнкт на категоризиращото име.

39.а. *Новият филиал* [A=(местоположение) в Кърджали] „Любен Каравелов“.

39.б. „Любен Каравелов“ е филиал и (той) е в Кърджали. [В Кърджали] е адюнкт.

39.в. *Новият филиал „Любен Каравелов“ в Кърджали*. [В Кърджали] е адюнкт на собственото име.

Модификаторът *област на приложение* е комплемент и той предхожда останалите задпоставени модификатори – адюнкти. Адюнктите на собственото име могат да се разполагат пред и след него (*фирма ЛИ 1 ЕООД на Лили Иванова, филиал „Любен Каравелов“ в Кърджали*).

5.5.10. Семантичен модел X

(а) Клас на съществителното собствено име: име на организация – съставна дума с вътрешна категоризираща дума.

(б) Семантичен клас на категоризиращото нарицателно име: название на организация (църковна организация, правителствена организация / институция, партия, компания, търговски обект, обществен обект, медийна организация, неправителствена организация, изпълнителска и творческа организация).

(в) Категоризиращото нарицателно име е вътрешно по отношение на собственото име.

(г) Семантичен модел:

C=(спецификатор)* – категоризиращо нарицателно име – K=(област на приложение)? – A=(притежател)? – (A=(принадлежност към обект) – A=(местоположение))?) – A=(спецификация)* – собствено име

(д) Примери, които илюстрират семантичния модел:

40.а. *Националното училище за танцово изкуство* [А=(местоположение) в *София*].

40.б. *Националното училище за танцово изкуство е училище и (то) е в София*. [*В София*] е адюнкт на собственото име.

Когато категоризиращото нарицателно име е станало част от собственото име – название на организация, се наблюдават следните възможности:

– Допустими модификатори на категоризиращото нарицателно име се реализират в самото собствено име (*Българска асоциация за компютърна лингвистика*);

– Допустими модификатори на категоризиращото нарицателно име, които не са реализирани в собственото име, се разполагат, както следва: спецификаторите преди съставното собствено име, а комплементите и адюнктите – след него (*известният Нов български университет в София*).

– Всички реализирани модификатори са на собственото име.

Ако има употребено външно категоризиращо име, то може да приеме собствени модификатори, отговарящи на семантичния модел.

41.а. *Сдружението* [К=(област на приложение) за *подпомагане на българските деца в чужбина*] „*Фондация за България*“.

41.б. „*Фондация за България*“ е *сдружение и (тя) е за подпомагане на българските деца в чужбина*. [*За подпомагане на българските деца в чужбина*] е комплемент на категоризиращото име.

42.а. *Сдружение* [А=(принадлежност към обект) на *регионалните дружества „Асоциация за европейско развитие“*].

42.б. „*Асоциация за европейско развитие*“ е *сдружение и (то) е на регионалните дружества*. [*На регионалните дружества*] е адюнкт на категоризиращото име.

43.а. *Сдружението* [А=(притежател) на *производителите на млечни продукти*] *Асоциация на млекопреработвателите в България*.

43.б. *Асоциация на млекопреработвателите в България е сдружение и (тя) е на производителите на млечни продукти*. [*На производителите на млечни продукти*] е адюнкт.

44.а. *Сдружение Асоциация на млекопреработвателите в България на производителите на млечни продукти*. [*На производителите на млечни продукти*] е адюнкт на собственото име.

45.а. *Сдружението* [А=(местоположение) в *Кърджали*] *Асоциация на млекопреработвателите в България*.

45.б. *Асоциация на млекопреработвателите в България е сдружение и (тя) е в Кърджали*. [*В Кърджали*] е адюнкт.

45.в. *Сдружение Асоциация на млекопреработвателите в България в Кърджали*. [*В Кърджали*] е адюнкт на собственото име.

Модификаторът *област на приложение* е комплемент и той предхожда останалите задпоставени модификатори – адюнкти. Адюнктите на собственото име могат да се разполагат и след него, в такъв случай категоризиращото име не се членува.

6. Синтактични структури

Езиково специфичните синтактични структури са именни фрази, чрез които се изразява даден семантичен модел. Синтактичните структури показват съответствието със семантичните структури – т.е. кой елемент от даден семантичен модел на кой елемент от дадена синтактична структура съответства. Синтактичните структури включват информация за част на речта на опората и конституентите; зависимостта между конституентите (съгласуване), ако е приложимо; словоредата и допълнителните зависимости при появата на конституентите; позицията на клитиките (притежателни местоименни и въпросителни клитики), когато са допустими; възможните обособявания, възможните синтактични алтернативи. Синтактичните структури са езиково специфични и включват комбинации от прилагателни имена във функцията на определения и в позиция преди името; местоимения във функция на определения в позиция преди името (притежателни и показателни местоимения); съществителни имена в позиция преди името във функция на приложения, местоименни клитики или предложни фрази в позиция след името и във функция на несъгласувани определения. Една или повече синтактични структури могат да изразяват един и същ семантичен модел. Например на семантичен модел I отговаря синтактична структура I (изключени са алтернативите, обособяването, допустимите съчинителни връзки и допустимите употреби на повече от една категоризираща дума), представена в таблица 1 – всеки ред е алтернативна възможност за реализация.

В синтаксиса нарицателното съществително (категоризиращата дума), което означава класа на съществителното собствено, се нарича приложение. В словосъчетанието със собствено име собственото име е главна част, а приложението – зависима. Казва се, че и двете съществителни в словосъчетания като *професор Пенчев* назовават един и същ референт и че са синтактично синонимни с копулативното изречение *Пенчев е професор*, в което главната част на словосъчетанието е трансформирана като подлог на изречението (Пенчев/Penchev 1984). Друго доказателство в подкрепа на приемането, че главна част на словосъчетания с роднински названия, професии, титли, звания, социални обекти (най-близо до организации), географски обекти е собственото име, привежда Алмалех (Almaleh 1993): обичайният ред на определенията в български е да са предпоставени, а приложенията са вид определения и също следват този ред. Наред с тези основания можем да посочим като основен аргумент за това, че съществителното собствено е главна част в

Таблица 1
Синтактична структура I

членувано прилагателно; (притежателна клитика)?	прилагателно*	категоризиращо съществително	предложна група (област на приложение)?	предложна група ((принадлежност към институция) (местоположение) (произход))?	предложна група (спецификация)?	собствено име
членувано притежателно метоимение	прилагателно*	категоризиращо съществително	предложна група (област на приложение)? предложна група (притежател)?	предложна група ((принадлежност към институция) (местоположение) (произход))?	предложна група (спецификация)?	собствено име
членувано прилагателно	прилагателно*	категоризиращо съществително	предложна група (област на приложение)? предложна група (притежател)?	предложна група ((принадлежност към институция) (местоположение) (произход))?	предложна група (спецификация)?	собствено име
		членувано категоризиращо съществително; (притежателна клитика)?	предложна група (област на приложение)?	предложна група ((принадлежност към институция) (местоположение) (произход))?	предложна група (спецификация)?	собствено име
		категоризиращо съществително				собствено име

словосъчетанията с приложение, факта, че собствените имена назовават конкретни лица, организации, географски обекти, които от своя страна могат да се характеризират като принадлежащи към различни по-големи или по-малки класове, изразявани от категоризиращото име – приложение. Семантичните основания намират потвърждение и в синтаксиса – ако собственото име може да се интерпретира еднозначно, то винаги може да се употреби самостоятелно, без категоризиращата дума, а е известно, че подчинителните словосъчетания се субституират от главната си част. В *Грамматика на съвременния български книжовен език. Том 3. Синтаксис* (Грамматика/Gramatika 1983, 182) приложенията, които означават названия на организации, заводи, институти, учреждения, вестници, списания, улици, квартали, книги и др., се приемат за главна част, а съществителните собствени – за тяхна подчинена част. Не смятаме, че има семантична или синтактична разлика между собствено име на лице и собствено име на организация или географски обект, така че подобно разграничение остава необосновано. Разлика съществува в степента на еднозначно асоцииране на името с назоваваното от него лице или обект и по тази причина категоризиращото име често е задължително за правилната интерпретация на географските обекти и организациите (площад „Славейков“, читалище „П. Р. Славейков“, улица „П. Р. Славейков“).

Ако няма модификатори, категоризиращото нарицателно име е неопределено (*министър Иванов, инженер Иванов, майор Иванов, но новият министър на образованието Иван Иванов*). Ако има модификатори, категоризиращото име е винаги определено и определеността се изразява или с определителен член, (рядко) с показателно местоимение (*известният професор Иванов, професорът ни по физика Иванов, тази баба Мария*) или (още рядко) с притежателна клитика (*баба ми Мария*). Неизменяемите предпоставени съществителни в начална позиция, показващи *област на употреба*, не се членуват, а се членува категоризиращото име (*ски учителят Иванов*). При обръщенията и някои титли на светии категоризиращото нарицателно име е винаги неопределено (*госпожо Петрова, свети Климент Охридски*). Ако категоризиращото нарицателно име е нечленувано, модификаторите на собственото име се разполагат след него (*професор Иванов от Софийския университет, архитект Иванов от Пловдив*).

Изказвано е мнението, че ако категоризиращото име е определено, то става главна част, а съществителното собствено – подчинена част (Андрейчин/Andreychin 1978; Пенчев/Penchev 1984; Алмалех/Almaleh 1993), т.е. в словосъчетанието *професорът по физика Иванов* главна част е категоризиращото име *професорът*. Според нас няма основание за промяна на главната част, тъй като няма промяна в изразяваната семантика: собственото име назовава конкретно лице или обект от действителността, а категоризиращото име посочва класа, към който може да се причисли лицето или обектът. Причината за задължителната поява на определителния член е появата на модификатори

към категоризиращото име. Правилата за съгласуване между съществителното собствено име, разширено с категоризиращо нарицателно име, и сказуемото в изречението не могат да се използват като проверка за главната част:

46.а. *Професорът по компютърна лингвистика Светла Коева е дошла в Пловдив.* (съгласуване с главната част – съществителното собствено)

46.б. **Професорът по компютърна лингвистика Светла Коева е дошъл в Пловдив.*

47.а. *Село Опълченец е разположено в Южна България.* (съгласуване с приложението)

47.б. * *Село Опълченец е разположен в Южна България.*

47.в. * *Опълченец е разположен в Южна България.*

48.а. *Стара Загора е разположена в Южна България.* (съгласуване с главната част – съществителното собствено)

48.б. *Град Стара Загора е разположен в Южна България.* (съгласуване с приложението)

48.в. * *Стара Загора е разположен в Южна България.*

Очевидно е, че при имената на организации и географските имена е възприето условно съгласуване с рода на категоризиращото нарицателно име, тъй като съществителното собствено може да не е достатъчно, за да се разграничи еднозначно семантиката на обозначавадения обект (може да съвпада със собствено име на друго лице, организация или географски обект или да съдържа друго категоризиращо име). Ако името, назоваващо организации и географски обекти, е достатъчно широко разпространено, за да може да се асоциира с обекта, който назовава, при самостоятелната му употреба съгласуването е с него (например *Варна е разположена*, но *Българи е разположено*).

Спецификаторите – числителни имена (рядко), притежателни местоимения (пълна форма), прилагателни имена (възможно е да има повече от едно), се явяват в препозиция по отношение на категоризиращото нарицателно име и се съгласуват по род и число. Възможно е да няма съгласуване със собственото име, ако то е от чужд произход или във форма, която не съвпада с рода и числото на категоризиращото име (при чужди фамилни имена за лица от женски пол, при географски имена и имена на организации) (*хубавата професор Смит, новата фирма „Съгласие“, китното село Българи*). Моделът за съгласуване е същият както при съгласуването със сказуемо: при личните имена съгласуването е в зависимост от пола на лицето, при имената на организации и географски обекти, ако името е употребено самостоятелно и обектът, който се назовава, е еднозначно различим, съгласуването е със собственото име (*хубавата Варна*), ако е употребено категоризиращо име – съгласуването е с него.

В структурата на словосъчетания с главна част лично име се срещат прилагателни, които семантично зависят от общия референт – лице (*хубавата професор Иванова*), и прилагателни, които зависят от специфичния референт – приложение (*асоциираният професор Иванова*), в такива случаи съгласу-

ването е с конкретната главна част. При географските имена и имената на организации прилагателните зависят от референта на приложението и съгласуването е с категоризиращото име.

Притежателните местоимения могат да се явяват в пълна и кратка форма. Ако се допуска модификатор – *притежател*, алтернативата е между притежателна предложна група в постпозиция, притежателно прилагателно в препозиция и притежателно местоимение в пълна или кратка форма (*бабата на Мая Мария, Маината баба Мария, нейната баба Мария, баба ѝ Мария*). Й. Пенчев (Penchev 1984, 78) посочва, че кратката форма на притежателното местоимение следва първата членувана дума в именното словосъчетание, като изключение се наблюдава при някои роднински имена в единствено число, които, употребени без предпоставени модификатори, могат да се съчетават с кратките форми на притежателните местоимения (*баба ми Мария*), но не могат – с притежателна предложна група в постпозиция (**баба на Даниел Мария*).

Комплементите – *област на употреба*, се изразяват със задпоставени предложни групи, които в някои случаи може да се редуват с неизменяеми предпоставени съществителни имена (*учителят по физика Иванов, инструкторът по тенис Иванов, тенис инструкторът Иванов*). Комплементите или адюнктите – *принадлежност към обект и местоположение*, се изразяват със задпоставени предложни групи. Адюнктите – *произход и спецификация*, се изразяват със задпоставени предложни групи. Предложната група за *произход* се редува с предпоставено прилагателно (*професорът от Лондон Смит, лондонският професор Смит*). Наблюдават се предложни групи, които могат да са към конкретния референт на собственото име (за *произход* при имена за лица; за *принадлежност към обект и местоположение* при имената на географски обекти с външно категоризиращо име; за *притежател* и *местоположение* при имената на организации с външно категоризиращо име) или към категоризиращото име – приложение (останалите при съответните класове собствени имена с външно категоризиращо име). При имената за географски обекти и организации, които съдържат вътрешно категоризиращо име, всички модификатори са към собственото име. Предложните групи, които са модификатори директно на собственото име, може да се разполагат както пред него, така и след него.

Въпросителната частица *ли* може да се появи след първия членуван модификатор, ако не се следва от кратка форма на притежателно местоимение или модификатор – *притежател*, или след цялото словосъчетание (*новият ли професор по физика Иванов, професорът по физика Иванов ли, професор Иванов от Пловдив ли, *професорът ни ли по физика Иванов, селото на Мария Опълченец ли, *селото на Мария ли Опълченец*).

Категоризиращите нарицателни имена се разполагат вляво от съществителното собствено. Възможно е два вида обособяване – без промяна на словоредата и с промяна на словоредата (Пенчев/Penchev 1984).

- 49.а. *Професорът по физика от Софийския университет Иван Иванов.*
49.б. *Професорът по физика от Софийския университет, Иван Иванов.*
49.в. *Иван Иванов, професор по физика от Софийския университет.*
49.г. *Иван Иванов, професорът по физика от Софийския университет.*

Първото обособяване (49.б.) може да се причисли към така нареченото рестриктивно обособяване, при което изпускането на обособената част *Иван Иванов* може да доведе до неразбиране на значението, ако има повече от един професори по физика от Софийския университет. Всяка една от разгледаните семантични структури може да се изрази както със словосъчетание без обособяване, така и със словосъчетание с обособяване. Изключение правят официалните обръщения (*господин*) и титлите на светии (*свети*), при които не може да има обособяване.

7. Заключение

Студията разглежда названията на лица, географски обекти и организации и предлага семантична и структурна класификация, мотивирани от семантичната и синтактичната структура на имената, обединени в отделен клас; описание на семантичните модели, с които се характеризират класовете имена на лица, организации и географски обекти, и общо представяне на особеностите на синтактичните структури, които кореспондират със семантичните модели.

Семантичните модели показват възможните комбинации между класове собствени имена (лично име; лично име, разширено с идентификатор; лично име, разширено с фамилно име; фамилно име; име на географски обект; име на географски обект – съставна дума с вътрешна категоризираща дума; име на организация; име на организация – съставна дума с вътрешна категоризираща дума) и семантичните класове на категоризиращите нарицателни имена (светски длъжности, професии и военни чинове; академични титли; аристократични титли и религиозни длъжности; титли на светци; роднински названия; обръщения; географски обекти; организации), както и семантиката на модификаторите: спецификатори, комплементи (*област на приложение, притежател, принадлежност към институция, местоположение*) и адюнкти (*произход, спецификация, притежател, принадлежност към обект, местоположение*). В синтактичните структури се описват детерминаторите, спецификаторите, комплементите и адюнктите – тяхното изразяване на синтактично равнище (част на речта, клитизация, съгласуване, словоред, възможности за синтактични алтернативи и обособяване). Представеният в студията модел на описание е приложен и за други езици – английски, гръцки, сръбски и френски (Коева и др. под печат/Koeva et al. in press).

Езиково независимите семантични модели и езиково специфичните синтактични структури, които им съответстват, могат да се използват при различ-

ни видове компютърна обработка на многоезиково съдържание, която изисква идентифициране на именувани обекти – имена на лица, на организации и на географски обекти и словосъчетанията с тях.

Литература

- Алмалех 1993:** Алмалех, М. *Семантика и синтаксис*. София: Университетско издателство „Св. Климент Охридски“, 1993.
- Андрейчин 1978:** Андрейчин Л. *Нова българска граматика*. 2. изд. София: Наука и изкуство, 1978.
- Граматика 1983:** *Граматика на съвременния български книжовен език. Том 3. Синтаксис*. София: Издателство на БАН, 1983.
- Пенчев 1984:** Пенчев, Й. *Строеж на българското изречение*. София: Наука и изкуство, 1984.
- Стоянова 2012:** Стоянова, Ив. *Автоматично разпознаване и тагиране на съставни лексикални единици в българския език*. Докторска дисертация. София: Институт за български език, 2012.
- Тодорова 2010:** Тодорова, М. *Съставни единици в Българския семантично аотиран корпус*. – В: *Българският семантично аотиран корпус*. София: Институт за български език, 2010, 166–185.

References

- Almaleh 1993:** Almaleh, M. *Semantika i sintaksis* [Semantics and Syntax]. Sofia: Universitetsko izdatelstvo „Sv. Kliment Ohridski“, 1993.
- Anderson 2007:** Anderson, J. *On the Grammar of Names*. Oxford: Oxford University Press, 2007.
- Andreychin 1978:** Andreychin, L. *Nova balgarska gramatika* [New Bulgarian Grammar]. Sofia: Nauka i izkustvo, 1978.
- Baldwin et al. 2003:** Baldwin, T., C. Bannard, T. Tanaka, D. Widdows. An empirical model of multiword expression decomposability. – In: *Proceedings of the ACL Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*. Sapporo, 2003, 89–96.
- Brunstein 2002:** Brunstein, A. Annotation guidelines for answer types. LDC2005T33, Linguistic Data Consortium, Philadelphia; catalog.ldc.upenn.edu/docs/LDC2005T33/BBN-Types-Subtypes.html
- Carroll 1985:** Carroll, J. *What's in a Name?* New York: W. H. Freeman, 1985.
- Culicover, Jackendoff 2005:** Culicover, P. W., R. Jackendoff. *Simpler Syntax*. Oxford: Oxford University Press, 2005.
- Ferrés et al. 2004:** Ferrés, D., M. Massot, M. Padró, H. Rodriguez, J. Turmo. Automatic classification of geographical named entities. – In: *Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004*, 1–4.
- Fleischman, Hovy 2002:** Fleischman, M., E. Hovy. Fine grained classification of named entities. – In *Proceedings of the 19th International Conference on Computational Linguistics-Volume 1*. Association for Computational Linguistics, 2002, 1–7.
- Giuliano 2009:** Giuliano, C. Fine-grained classification of named entities exploiting latent semantic kernels. – In: *Proceedings of the Thirteenth Conference on Computational Natural Language Learning (CoNLL'09)*. Stroudsburg, PA: ACL, 2009, 201–209.
- Gramatika 1983:** *Gramatika na savremenniya balgarski knizhoven ezik. Tom 3. Sintaksis*. [Grammar of the Modern Bulgarian Language. Vol. 3. Syntax]. Sofia: Izdatelstvo na BAN, 1983.

- Hwang 2011:** Hwang, J. D. *Making Verb Argument Adjunct Distinctions in English. Synthesis Paper*. Boulder, Colorado: University of Colorado.
- Huddleston 1988:** Huddleston, R. *English Grammar: an Outline*. Cambridge: Cambridge University Press, 1988.
- Kamp, Reyle 1993:** Kamp, H., U. Reyle. From Discourse to Logic: an Introduction to Model Theoretic Semantics – In: *Natural Language, Formal Logic and Discourse Representation Theory*. Vol. 1. Dordrecht: Kluwer Academic Publishers, 1993.
- Koeva, Dimitrova 2015:** Koeva, S., T. Dimitrova. Rule-based person named entity recognition for Bulgarian. Slavic languages in the perspective of formal grammar. – In: *Proceedings of FDSL 10.5, Brno 2014* (Series Linguistik International, vol. 37). Bern: Peter Lang, 2015, 121–137.
- Koeva et al. in press:** Koeva, S., C. Krstev, D. Vitas, T. Kyriacopoulou, C. Martineau, T. Dimitrova. Semantic and syntactic patterns of multiword names (a cross-language study). – In: *Multiword Expressions: Insights from a Multilingual Perspective*. Berlin: Language Science Press (in press).
- Kripke 1980:** Kripke, S. *Naming and Necessity*. Rev. ed. Cambridge, MA: Harvard University Press, 1980.
- Leseva, Stoyanova 2007:** Leseva, S., I. Stoyanova. Treatment of named entities in machine translation. – In: *Proceedings of the 2007 International Nooj Conference*. Cambridge: Cambridge Scholars Publishing, 2007, 254–272.
- Magnini et al. 2002:** Magnini, B., M. Negri, R. Prevete, H. Tanev. A wordnet-based approach to named entities recognition. – In: *Proceedings of the 2002 Workshop on Building and Using Semantic Networks*. Vol. 11. Stroudsburg, PA: ACL, 2002, 1–7.
- McDonald 1996:** McDonald, D. D. Internal and external evidence in the identification and semantic categorisation of proper names. – In: *Corpus Processing for Lexical Acquisition*. Cambridge: MIT Press, 1996, 32–43.
- Nadeau, Sekine 2007:** Nadeau, D., S. Sekine. A survey of named entity recognition and classification. – *Linguisticae Investigationes*, 2007, 30, 1, 3–26.
- Osenova, Kolkovska 2002:** Osenova, P., S. Kolkovska. Combining the named-entity recognition task and NP chunking strategy for robust pre-processing. – In: *Proceedings of The First Workshop on Treebanks and Linguistic Theories (TLT2002), 20th and 21st September 2002, Sozopol, Bulgaria, 2002*, 167–182.
- Payne 2010:** Payne, T. E. Noun Complements vs. Post-Nominal Modifiers: Additional Text for Chapter Nine. – In: *Understanding English Grammar*. Cambridge: Cambridge University Press; pages.uoregon.edu/tpayne/UEG/UEG-additionalreading-ch9-nouncomplements.pdf
- Payne 2011:** Payne, T. E. *Understanding English Grammar*. Cambridge: Cambridge University Press, 2011.
- Penchev 1984:** Penchev, J. *Stroezh na balgarskoto izrechenie* [Structure of the Bulgarian Sentence]. Sofia: Nauka i izkustvo, 1984.
- Putnam 1975:** Putnam, H. *Mind, Language and Reality*. New York: Cambridge University Press, 1975.
- Radford 1988:** Andrew Radford. *Transformational Grammar: A First Course*. Cambridge: Cambridge University Press, 1988.
- Rodríguez et al. 1998:** Rodríguez, H., S. Climent, P. Vossen, L. Bloksma, W. Peters, A. Alonge, F. Bertagna, A. Roventini. The top-down strategy for building EuroWordNet: Vocabulary coverage, base concepts and top ontology. – In: *EuroWordNet: A Multilingual Database with Lexical Semantic Networks*. Dordrecht: Kluwer Academic Publishers, 1998, 45–80.
- Searle 1958:** Searle, J. R. Proper names. – *Mind*, 1958, 67.266, 166–173.

- Sekine, Nobata 2004:** Sekine, S., C. Nobata. Definition, dictionaries and tagger for Extended Named Entity Hierarchy. – In: *Proceedings of LREC'04*. ELRC, 2004, 1977–1980.
- Stoyanova 2012:** Stoyanova, I. *Avtomatichno razpoznavane i tagirane na sastavni leksikalni edinitsi v balgarskiya ezik* [Automatic Identification and Tagging of Multiword Expressions in Bulgarian]. Doktorska disertatsiya [Doctoral dissertation]. Sofia: Institut za balgarski ezik, 2012.
- Stoyanova 2014:** Stoyanova, I. Automatic categorisation of multiword expressions and named entities in Bulgarian. – In: *Proceedings of CLIB-2014*, Sofia: Institut za balgarski ezik, 2014, 40–48.
- Todorova 2010:** Todorova, M. Sastavni edinitsi v Balgarskiya semantichno anotiran korpus [Multiword expressions in the Bulgarian Sense-Annotated Corpus]. – In: *Balgarskiyat semantichno anotiran korpus* [Bulgarian Semantically Annotated Corpus]. Sofia: Institut za balgarski ezik, 2010, 166–185.