

СУ „Св. Климент Охридски”
Факултет по славянски филологии



МП „Компютърна лингвистика. Интернет
технологии в хуманитаристиката”

Магистърска теза на тема:
Български онлайн лексикон
(базиран на ФреймНет)

Студент:

Вилислав Василев
771141

Научен ръководител:

доц. д-р Атанас
Атанасов

София,
2019

Съдържание:

1.	Увод.....
2.	Въведение в проблематиката.....
3.	Теоретична рамка на изследването.....
3.1.	Цели и задачи.....
3.2.	ФреймНет – значение и особености.....
3.3.	Българският ФреймНет.....
3.4.	Немският ФреймНет.....
4.	Структура на българския онлайн лексикон (БОЛ).....
4.1.	Компютърните езици XML и DTD.....
4.2.	Описание на XML-кода.....
4.2.1.	Елемент <framenet_description>.....
4.2.2.	Елемент <frameelements>.....
4.2.3.	Елемент <words>.....
4.2.3.1.	Елемент <expr>.....
4.2.3.2.	Елемент <meaning>.....
4.2.3.3.	Елемент <grammar>.....
4.2.3.4.	Елемент <example_sentences>.....
4.2.3.5.	Елемент <templates>.....
4.2.3.6.	Елемент <other_forms>.....
4.3.	Лингвистична информация.....
4.3.1.	Граматични правила.....
4.3.2.	Първа група.....
4.3.3.	Втора група.....
5.	Обобщение на резултатите.....
5.1.	Особености при елементите <expr>.....
5.2.	Особености при елементите <grammar>.....
5.3.	Особености при елементите <example_gr> и <template_gr>.....
6.	Заклучение.....
7.	Приложения.....
7.1.	Приложение 1.....
7.2.	Приложение 2.....
8.	Библиография.....

1. Увод

В областта на компютърната лингвистика и машинната обработка на естествените езици има множество изследвания по темата за т.нар. Фреймова граматика на Ч. Филмор (Филмор 1982), за самите фреймове и техните задължителни и допустими обкръжения. В световен мащаб може да се посочи английският ФреймНет (Бейкър и др. 1998, Филмор 2002), който е служил за основа на много следващи наблюдения. По негов аналог са създадени допълнителни електронни ресурси за испански, немски, японски, китайски и др. На български език също има подобен труд, въпреки че се откриват известни разлики между него и английския еквивалент. Семантико-синтактичният речник, по-известен като българския ФреймНет, разглежда проблематиката на българските глаголи на синтактично и семантично равнище (Коева и др. 2008).

Настоящото изследване не трябва да се разглежда като алтернатива и конкурент, а като негово допълнение и обогатяване, тъй като освен глаголите то обхваща и други части на речта. Електронният ресурс е създаден по подобие на немския ФреймНет (<http://coerll.utexas.edu/frames/>) и затова може да се интерпретира като електронен речник и онлайн лексикон. Това, което го различава от българския ФреймНет, е намерението да служи като един вид паралелен корпус, защото лексиконът засяга един конкретен фрейм, който е разгледан и анализиран на три езика: английски, немски и български. Проследява се точно какви лексикални единици се извикват, привежда се съответната дефиниция от БулНет (българския WordNet), при необходимост се добавя информация за конкретни граматични правила. За всяка лексикална единица са добавени примери, с които се илюстрират дадените прилики и разлики между трите езика. На този етап може да се направи разлика между задължителни и допустими елементи в обкръжението на съответната лема. На следваща стъпка се представят синтактичните особености посредством темплейти, които задават броя на аргументите и адюнктите. Към края на всяка лексикална единица са въведени също така допълнителни граматични форми, напр. глаголни времена, множествено число и т.н., но те не са обект на изследването, а имат по-скоро илюстративен характер.

Лексиконът може да се използва както в академични среди за по-задълбочени наблюдения, така и от студенти и ученици при изучаване и сравняване на някой от представените езици. По подобие на първия ФреймНет и настоящият речник е създаден с помощта на компютърния език XML (Extensible Mark-up Language), който с времето се е доказал като един от най-надеждните и стабилни езици за кодиране на

онлайн речници. Чрез DTD (Document Type Definition) са направени дефиниции за всеки елемент и атрибут от XML-кода, а визуализацията – посредством eXist-db сървър, на който е инсталиран проектът, подобно на системата SynTags (Атанасов 2016).

2. Въведение в проблематиката

Изследването представя един нов вид педагогически речник, който се различава от традиционните хартиени и електронни речници, които съдържат само дефиниции и принадлежност към определена част на речта, но не показват как точно могат да се прилагат съответните думи в разговорния език. По какво се откроява от широко разпространените ресурси? Могат да се обособят три основни разлики:

- при него лемите са организирани според своите значения, а не по азбучен ред;
- предоставя подробни дефиниции и посочва специфични отличителни белези между привидно еднакви по значение думи на немски, английски и български език;
- включва редица реални примери за всяка дума, които са заети от немския ФреймНет и са преведени на български език за подробна илюстрация.

Електронният лексикон (<http://coerll.utexas.edu/frames/>) се базира на немския ФреймНет в Тексаския университет в Остин, а той представлява цифров архив за това как се използват немски думи в реалния живот. Предоставя се актуална база от данни за синтактичните, семантичните и прагматичните контексти, в които се появяват лемите. Тя е организирана под формата на систематично подредени думи, детайлното им описание и определяне на контекста, в който се използват. Освен това се набляга върху разликите при употребата на немски, английски и български фреймове. Всичко това е подкрепено от конкретни примери и граматични данни, в случай че такива са необходими. Всеки лексикален запис съдържа информация за регистъра на дадената дума, честотата и свързаните с нея значения. Посредством веб-базираната структура ресурсът може да се актуализира с нови думи или нови употреби на съществуващи думи.

За да се разбере цялостната идея на ФреймНет-базираните ресурси, е необходимо да се изясни, че за всяка една ситуация от действителността може да се причисли точно определено събитие, действие и участници. От това становище се прави и диференциацията на синтактично равнище между аргументи и адюнкти. Аргументите са задължителните участниците в ситуацията, свързани с определени

семантични отношения, а адюнктите са незадължителни участници. Тази дефиниция е от значение за разбирането на фреймовете и проследяването на разлики, когато започва описанието в следващите глави на изследването.

3. Теоретична рамка на изследването

3.1. Цели и задачи

Настоящата разработка може да се разгледа като начален етап от едно по-задълбочено изследване, свързано със семантичните фреймове и тяхното обкръжение. В тази фаза се анализира и сравнява само един конкретен фрейм през погледа на три езика. Тъй като кодът на речника е отворен, може да бъдат направени и допълнителни наблюдения за други фреймове.

Идеята е да се помогне на ученици и студенти при изучаването и употребата на думите в съвременния немски, английски и български език. Направен е опит за обстойно описание на съчетаемостта между отделните части на речта и техните задължителни и допустими елементи чрез създаването на платформа за сравняване на германски със славянски езици. Цялата информация се представя на немски, английски и български, като се преследва стремежа за направата на достатъчно голям електронен ресурс, който да съдържа конкретни данни за ограниченията при синтактичната и семантичната съчетаемост на съответните единици.

Многоезиковият паралелен ФреймНет трябва да съдържа достатъчно анотирани примери, за да е в състояние да служи като помощно средство при евентуално проследяване на актуалното състояние на съвременния език, съответно за немски, английски и български. За изпълнението на тази задача най-подходящ ресурс в съвременния български език е БулНет (<http://dcl.bas.bg/bulnet/>), тъй като в него думите са групирани в синонимни гнезда с обща дефиниция. Благодарение на тези синонимни множества (Synsets) БулНет спомага за по-лесното разпознаване и съотнасяне на възможните разширения при отделните лексикални единици за разлика от традиционните тълковни речници и други подобни ресурси. Чрез тях се откриват множество подчинени елементи, които също са приведени като част от изследването. От друга страна, в БулНет са включени всички основни понятия от WordNet, което позволява да бъдат избрани най-подходящите синсети за съответните единици (Атанасов 2016).

Конкретната цел е да се създаде електронен ресурс под формата на онлайн лексикон, който да обхваща голям брой от анотирани лексикални единици на трите езика, принадлежащи към семантичния фрейм *die Arbeit/work/работа*, и да се

изгради многоезиков и паралелен корпус. За да се изпълни тази задача, в електронния речник са представени преводи на съответните английски и немски примери. Още няколко задачи биха могли да се дефинират за постигане на поставената цел:

- откриване на възможно най-много лексикални единици, които биват извиквани на български език при сравнение на примери от съответните английски и немски варианти;
- аотиране на всички примери на немски, английски и български език, като се анализира какви фреймови елементи се извикват от съответната лексикална единица;
- определяне на съчетаемостта – конкретни данни за аргументи и адюнкти;
- откриване на прилики и разлики сред трите езика и анализ на получените резултати.

Като основа на настоящото изследване са приети Фреймовата граматика на Филмор (Филмор 1982), моделите на английския ФреймНет (Бейкър и др. 1998, Филмор 2002), както и електронният формат на немския ФреймНет (<http://coerll.utexas.edu/frames/>), който също е създаден въз основа на английските концепции.

3.2. ФреймНет – значение и особености

Преди настоящото изследване да пристъпи към по-задълбочени теоретични и лингвистични наблюдения, е важно да се опише какво представлява ФреймНет и с какви специфики се характеризира. Целта на това начинание е да се добие предварителна представа за проекта и начина на работа, за да може да се разбере впоследствие как функционира българският онлайн лексикон.

В лингвистиката съществува голямо количество литература и електронни ресурси, които описват семантично ситуации от действителността. Ще бъдат въведени само тези, които оказват по-сериозно влияние върху настоящото наблюдение. Както бе описано в уводната част са създадени ресурси на множество езици. Могат да бъдат различени два основни подхода (Коева и др. 2008). От една страна, те са ориентирани към описание на фреймове. Към тях се числят повече от една лексикална единица. От друга страна, може да се обособи такъв тип, който описва конкретни лексикални единици и техните обкръжения.

При дефинирането на ФреймНет като понятие разработката се основава на английския ФреймНет, защото той е послужил като основа на всички последвали изследвания. Проектът ФреймНет е разработен от университета в Бъркли и

представлява лексикална база от данни, която отразява семантичната и синтактичната съчетаемост между думите в реална ситуация. Изследванията на Чарлз Филмор (Филмор 1982) за теорията на значенията лежат в основата на проекта. Т.нар. Фреймова семантика помага за разбирането на значението на думите посредством семантични фреймове. Семантичен фрейм е концептуална структура, която кодира културно знание за дадена ситуация и нейните участници (Коева и др. 2008). Те са елементи, които характеризират поведението на участниците в ситуацията. За да може човек да осъзнае значението на думите, трябва да добие информация за техните структури, които всъщност представляват условие за съществуването на думите в естествените езици и тяхната употреба.

Към днешна дата проектът ФреймНет наброява 1224 фрейма, които съдържат 13640 лексикални единици. Краткият формален анализ на английския ФреймНет в разработката се основава на изследванията на А. Атанасов по проблематиката на предложните фрази (Атанасов 2016), където във всеки фрейм се отличават три нива:

- фреймови елементи (напр. работа);
- синтактична функция (напр. подлог);
- типа на фразата (напр. РР).

Всяка лексикална единица се разглежда като свързана двойка от дума и значение. Фреймовете описват определени ситуации от реалния живот. За по-голяма яснота ще бъде посочен фреймът Прилагам_топлина (Apply_heat). За него са характерни следните фреймови елементи: ГОТВАЧ (COOK); ХРАНА (FOOD) и НАГРАВЯЩ_ИНСТРУМЕНТ (HEATING_INSTRUMENT).

При изследването на Св. Коева дори е направена по-задълбочена диференциация на елементите (Коева и др. 2008). Тук те се разделят на:

- ядрени елементи – елементи, които са задължителни и правят фрейма неповторим;
- периферни елементи – елементи, които не характеризират фрейма уникално и могат да се срещнат и при други фреймове (Време, Място, Начин и т.н.);
- екстратегматични елементи – елементи, които не са задължителни и имат по-скоро абстрактен характер.

Фреймовите елементи могат да бъдат извиквани от думи като пека, варя и т.н. В този случай точно тези думи, които извикват фрейма, се наричат лексикални единици. Всяка единица съдържа ръчно анотирани примери, които за целия проект са малко над 174000. В повечето случаи глаголите изпълняват ролята на извикващи думи, а фреймовите елементи са негови аргументи. Отглаголни съществителни,

причастия и в по-малка степен съществителни имена също могат да заемат тази роля.

3.3. Българският ФреймНет

Въпреки че в началото на разработката беше описано с няколко думи какво представлява българския ФреймНет и какви особености има, в тази глава ще бъде отделено по-специално внимание на неговата структура. На този етап от проекта са аотирани задължителните обкръжения на българските глаголи в синтактични фреймове на синтактично и лексикално-семантично равнище. Както посочва А. Атанасов (Атанасов 2016) концептуалният модел на българския ФреймНет обхваща допълнителни типове данни, като по този начин се надгражда информацията от английския еквивалент. Към дефиницията, лексикално-семантичната информация и аотирани примери се добавя морфологична информация.

Системата на проекта включва лексикална единица, семантичен фрейм, граматичен клас и синтактичен фрейм. Всички те образуват речникова единица във ФреймНет (Коева и др. 2008). Всяка лексикална единица (глагол) изразява уникално лексикално значение и се състои от:

- лема

Към момента обект на изследване в проекта са само глаголите.

- тълковна дефиниция

За дефинициите авторите на проекта са се позовали на вижданията на Ландау (Ландау 1989), който изтъква редица принципи, които трябва да се спазват при дефинирането на значенията. Основното изискване е дефинициите да описват и интерпретират уникално тълковно понятие със средствата на езика. За целта се използват дефинициите от БулНет, защото при него думите са организирани в синонимни множества и тази организация спомага за осъществяването на връзки със синонимни множества от други езици. За да се опише подробно всеки глагол, създателите на семантико-синтактичния речник въвеждат допълнителното изискване за експлицитно описание на лексикализацията на аргументите (Коева и др. 2008).

- аотирани примери

Приложени са многобройни примери, които маркират глагола и неговите задължителни и незадължителни обкръжения. Чрез тях се илюстрира съчетаемостта при отделните единици.

При семантичния фрейм всяка лексикална единица със съответното уникално значение прави връзка с подходящия фрейм от английския ФреймНет. С фрейма Рязане (Cutting) се илюстрира тази релация: Агентивен обект реже Цял обект на

Части, като използва Инструмент (An Agent cuts an Item into Pieces using an Instrument). На този етап става видна диференциацията между ядрени и периферни елементи.

Граматичният клас се характеризира с категориите парадигматичност, инхерентна транзитивност, вид на глагола и флективен тип, като по този начин се специфицира множеството от морфо-синтактични свойства на лемата (Коева и др. 2008).

- парадигматичност

Тук всеки глагол се категоризира като личен, безличен и третоличен.

- инхерентна транзитивност

Всеки глагол се дефинира като транзитивен и нетранзитивен. В рамките на тези групи съществуват допълнителни класове в зависимост от лексикалните свойства.

- вид на глагола

Лексикално-граматичната категория вид изразява начина, по който протича действието. Свършеният вид представя действието в неговата цялостност и комплексност, а несвършеният вид – в процеса на протичането му.

- флективен тип

Трите упоменати компонента определят граматичния клас в речника. Семантико-синтактичният речник черпи информация за флективните типове от Граматичния речник (Коева 1998а). С негова помощ се генерират всички валидни форми на глаголите.

Последната част на българския ФреймНет е синтактичният фрейм. При описанието на модела създателите на проекта са се позовали на основните трудове по български синтаксис (Пенчев 1984, 1993; Бъркалова 1997; Коева 1999, 2001, 2005б и други). Той съдържа една или повече синтактични структури. Тези структури определят броя на аргументите чрез стойностите на категориите:

- синтактична категория

За тази категория са възможни следните реализации под формата на синтактични фрази: NP, PP, AdvP, AP, S, SC/AP, SC/NP, SC/PP, CL/NP или CL/PP, NP/AdvP, PP/AdvP. В българския ФреймНет се разглеждат и анализират само синтактичните реализации на задължителните участници, т.е. само ядрените елементи от фреймовете са обект на изследването.

- експлицитност

На синтактично равнище аргументите се различават по начина на реализация, която е експлицитна или имплицитна. В речника фразите могат да бъдат задължително експлицитни или да останат имплицитни.

- предлози

В проекта предлозите се разглеждат като компонент, който лексикализира дадена семантична връзка между предикат и аргумент, между предикат и два аргумента и между предикат и три аргумента. Като цяло предлозите са зависими от семантиката на глагола. В българския ФреймНет предлозите с еднакво значение се числят към едно синонимно множество, аналогично на синонимните множества от БулНет.

- въвеждащи думи

Разглеждат се тези думи, които въвеждат подчинени изречения. Предмет на изследването са само допустимите въвеждащите думи към аргументни подчинени изречения. От това следва и разделението на аргументните изречения на субектни, допълнителни и обстоятелствени. Като стойности са дефинирани подчинителните съюзи *да* и *че*, въпросителна дума и обстоятелствен съюз.

- синтактични функции

Синтактичните функции са различни за отделните синтактични категории. Но авторите са направили общ списък за проекта: подлог, пряко и непряко допълнение, обстоятелствено пояснение, малко изречение, субектно изречение, допълнително изречение или обстоятелствено изречение.

- семантични ограничения

Семантичните ограничения се явяват последната част от синтактичния фрейм. Към тях спадат съществителните имена, независимо дали са въведени с или без предлог. В проекта те се класифицират в две основни групи. За селективните ограничения е характерна подялбата на съществителните имена на абстрактни и конкретни, одушевени и неодушевени, лица и нелица. За лексикално-семантичните ограничения авторите се доверяват на синонимния ред от българския WordNet, който доминира над множеството думи (Коева и др. 2008).

В края на тази глава може да се обобщи, че именно синтактичният фрейм в семантико-синтактичния речник е този уникален компонент, който го различава от английския ФреймНет и от онлайн лексикона, който ще бъде обект на изследване в предстоящите глави на разработката.

3.4. Немският ФреймНет

Немският ФреймНет (<http://www.laits.utexas.edu/gframenet/>) е създаден в Тексаския университет в Остин и цели съставянето на електронен лексикален ресурс за немски глаголи, съществителни и прилагателни имена. Въз основа на Фреймовата семантика (Филмор 1982) и корпуса с анотирани примери от съвременни текстове немският ФреймНет представлява пълен набор от семантични и синтактични

релации между отделни думи и техните значения. Получената база от данни се състои от лексикални записи, които съдържат информация за семантичните фреймове, към които принадлежат лексикалните единици, за задължителните и допустимите обкръжения на семантично и синтактично равнище, както и корпус от аотирани примери. Получените данни от проекта могат да се прилагат в областта на естествените езици и за чуждоезиково обучение.

Немският ФреймНет работи в сътрудничество с английския проект ФреймНет на Международния институт по компютърни науки в Бъркли, както и с проекта SALSA на университета в Саарланд. На неговата основа е създаден и немският онлайн лексикон (German Frame-semantic Online Lexicon), който всъщност е обект на изследване в настоящата разработка. Това е електронен ресурс от съвсем нов тип, който е качен на сървъра на Центъра за отворени образователни ресурси и езиково обучение към Тексаския университет в Остин и е достъпен на интернет страницата на проекта: <http://coerll.utexas.edu/frames/>.

В този речник немски леми и изрази са групирани според значението си в отделни фреймове, към които спадат всички думи (лексикални единици), свързани с тях на семантично равнище. За момента G-FOL се състои от 20 фрейма, сред които *arguing*, *cleaning*, *desire*, *exercise*, *sleep*, *thinking*, *work* и т.н. Те са с различна големина – част от тях съдържат по няколко фреймови елемента, докато други са с по-малък брой. Създателите на проекта онагледяват структурата на лексикона във видеоклип, който се намира в началото на сайта. С помощта на фрейма *Eating and Drinking* се прави кратък анализ на уебстраницата. За да се разбере фреймът като цяло, човек трябва да разполага с минимални познания за съответния език и да има опит с процеса приемане на храна. В този случай някакво живо същество поставя нещо в тялото си, за да го погълне. Участници в ситуацията са *Ingestor* (човек, който яде и пие) и *Ingestibles* (храна), които функционират като фреймови елементи (*frameelements/FE*) на фрейма *Eating and Drinking*. Поставяйки курсора върху съответния елемент автоматично се показва и неговата дефиниция в отделен фрагмент. В началото фреймът е описан с няколко думи в компонента *Frame description*. В случай че са необходими допълнителни обяснения и дефиниции, те са въведени в секцията *Details*, където също така се изтъкват разлики при употребата на немски и английски леми, ако такива са налични. Фреймовете са част от семантичната рамка на различни думи. Конкретният фрейм може да се извиква от глаголи, като например *essen/ to eat*, *trinken/ to drink*, *saufen/to booze* и др., от някои съществителни имена: *der Imbiss/snack* и т.н. Словосъчетания и други части на речта

също могат да извикат фрейма. Всички те спадат към категорията лексикални единици (*lexical unit/LU*).

В онлайн лексикона за обемните фреймове съществуват подчинени фреймове, които помагат за по-лесното класифициране на лексикалните единици, което всъщност го различава от класическите електронни ресурси. За вече разгледания пример *Eating and Drinking* това са:

- Eating – съдържа единици само за процеса хранене;
- Drinking – съдържа единици само за процеса пиене.

Те се отварят с помощта на падащи менюта, което улеснява потребителя при търсене на конкретна информация за дадена лексикална единица. Разбира се, те са обединени и в подчинен фрейм *Eating and Drinking*, в който са изброени всички възможни лексикални единици.

За всяка лексикална единица са добавени аотирани примери на немски език, за които е направен превод на английски, като създателите на лексикона са се стремели към буквален превод, за да се представят по-добре фреймовите елементи в изреченията. Фреймовите елементи от своя страна са оцветени в различни цветове, които са запазени в примерите. Тези изречения, които при буквален превод са неграматични, са променени по начин, който да отговаря на граматичните правила в английски. В случай че съществува разлика между немски и английски на равнище граматика, за съответната лексикална единица са добавени данни за граматичните правила в немски. В края на всяка лексикална единица има темплейти, които обозначават, от една страна, всички задължителни участници в ситуацията, и от друга, позицията им в изречението.

Посредством тази иновативна структура авторите на G-FOL имат намерението да представят на потребителите електронен ресурс от съвсем нов тип и да улеснят изучаващите езика.

4. Структура на българския онлайн лексикон (БОЛ)

4.1. Компютърните езици XML и DTD

Както вече бе упоменато българският онлайн лексикон анализира единствено фрейма *work/die Arbeit/работа* и е написан на компютърния език XML. Extensible Mark-up Language, или разширяем маркиращ език, е стандарт, който дефинира правила за създаване на специализирани маркиращи езици. Той указва как да бъде структуриран един документ чрез поставяне на определени етикети. Тези маркери от своя страна нямат собствено значение, а поясняват само граматиката

(синтаксиса) на езика. Ползата от XML се състои в това, че синтаксисът (структурирането) на документите се измисля веднъж, а специализираните маркиращи езици дефинират семантиката, т.е. набора от маркиращи етикети и тяхното значение. XML е създаден като олекотена версия на SGML и съответно е добре пригоден за употреба в интернет пространството.

В XML информацията се маркира посредством заграждане с етикети и по този начин се отбелязва, че маркираният къс информация трябва да бъде разглеждан като едно цяло, имащо определено значение. Етикетите вървят по двойки – отварящ и затварящ таг, посочвайки началото и края на обозначения текст. Освен това те могат да се влагат един в друг, така че да се получи йерархична организация на данните.

Това, което XML указва, е как да бъдат използвани етикетите. Всеки документ трябва да започва с коренов етикет, така че цялата информация да бъде маркирана от поне един етикет. Документ, който се подчинява на правилата, описани в XML-стандарта, се нарича добре оформен XML-документ. Документи, които отговарят на правилата, описани в дадена XML-схема, се наричат валидни. Валидните документи са винаги добре оформени. Една от най-простите XML схеми е DTD (Document Type Definition – дефиниция на документен тип). Дефиницията съдържа само декларации и правила и структурата не е йерархична. Чрез стринга ELEMENT се дефинират отделните етикети от XML-кода, а чрез ATTLIST – отделните атрибути под формата на атрибутен списък.

4.2. Описание на XML-кода

В тази глава ще се направи техническо описание на лексикона, т.е. ще бъде представена структурата на кода и всички негови спецификации. Той съдържа множество елементи, принадлежащи към тях дъщерни елементи и атрибути към част от елементите. Схемата, на която отговаря XML-документът, се посочва в началото му и следва определен от XML синтаксис. Първият ред от XML-документа се нарича още декларативен (декларация).

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<!DOCTYPE framenet SYSTEM "framenet.dtd">
```

Като ядрен елемент се обособява <framenet>, който съдържа елемента <frame>. Дъщерни елементи на <frame> са <framenet_description>, <frameelements> и <words>. Те от своя страна също съдържат дъщерни елементи с атрибути, на които ще се обърне повече внимание в следващите редове. В примера е представена основната структура на кода:

```
<framenet>
```

```

    <frame>
      <framenet_description>
        <...>
      </framenet_description>
      <frameelements>
        <...>
      </frameelements>
      <words>
        <...>
      </words>
    </frame>
  </framenet>

```

В DTD-файла елементите <framenet> и <frame> са съответно ядрен и дъщерен и имат следната дефиниция:

```

<!ELEMENT framenet (frame)>
<!ELEMENT frame (framenet_description, frameelements, words)>

```

4.2.1. Елемент <framenet_description>

Във <framenet_description> се прави описание на фрейма *work/die Arbeit/работа*. Като дъщерен елемент може да се посочи <definition>, който съдържа атрибут *lang*. Този атрибут има три различни стойности en | de | bg, които разграничават описанието на значенията на три различни езика.

```

<framenet_description>
  <definition lang="..."> ... </definition>
</framenet_description>

```

Елементите <framenet_description> и <definition> имат следната DTD-дефиниция, като при втория са застъпени вече анализирани атрибути, които също са отразени в дефиницията. На този етап се изразява отношението на принадлежност на подчинените елементи:

```

<!ELEMENT framenet_description (definition)+>
<!ELEMENT definition (#PCDATA| fe)*>
<!ATTLIST definition
  lang (en | de | bg) #REQUIRED
>

```

От тази кодировка става ясно, че <definition> има дъщерен елемент <fe> с допълнителни атрибути. Той също е широко разпространен в XML-кода, като чрез него се представят фреймовите елементи. Те са важна част от цялото изследване, защото са дъщерни елементи и за други етикети в изследването. В следващите редове ще бъдат разгледани детайлно.

```
<!ELEMENT fe (#PCDATA)>
<!ATTLIST fe
href (Employee | Employer | Position | Field | Task | Compensation) #REQUIRED
>
```

➤ английски език

За английски език е ползвана дефиницията от немския онлайн лексикон: These LUs describe various aspects of employment, where an Employee works for Employer in a certain Position or Field, completing assigned Tasks for Compensation. Both Employer and Employee must agree to enter into the employment relationship (the Employer hires the Employee, and the Employee agrees to fill the Position in exchange for Compensation), and either party can end the employment relationship (the Employee can quit, or the Employer can fire the Employee).

```
<definition lang="en">These LUs describe various aspects of employment, where
an Employee works for Employer ... </definition>
```

➤ немски език

За да се изпълни задачата на паралелен ФреймНет, в българския лексикон е въведена дефиниция и на немски език, която е: Diese lexikalische Einheiten beschreiben verschiedene Aspekte der Beschäftigung, bei denen ein Mitarbeiter für einen Arbeitgeber in einer bestimmten Position oder einem bestimmten Feld arbeitet und zugewiesene Aufgaben für die Vergütung abschließt. Sowohl Arbeitgeber als auch Arbeitnehmer müssen zustimmen, das Arbeitsverhältnis einzugehen (der Arbeitgeber stellt den Arbeitnehmer ein, und der Arbeitnehmer stimmt zu, die Position gegen Vergütung zu besetzen), und beide Parteien können das Arbeitsverhältnis beenden (der Arbeitnehmer kann kündigen oder der Arbeitgeber kann den Mitarbeiter entlassen).

```
<definition lang="de">Diese lexikalische Einheiten beschreiben verschiedene
Aspekte der Beschäftigung, bei denen ein Mitarbeiter ... </definition>
```

➤ български език

За елемента, свързан с български език, е направен превод на описанието на фрейма *work/die Arbeit/работа*: Тези лексикални единици описват различни аспекти на заетостта, при които Служителят работи за Работодателя на определена Позиция или Поле и изпълнява поставените Задачи за Възнаграждение. Работодателят и Служителят трябва да се съгласят да сключат трудовото правоотношение (Работодателят наема Служителя и Служителят се съгласява да заеме Позицията в замяна на обезщетение) и всяка от страните може да прекрати трудовото правоотношение (Служителят може да напусне или работодателят може да уволни Служителя).

```
<definition lang="bg">Тези лексикални единици описват различни аспекти на заетостта, при които Служителят работи за Работодателя ... </definition>
```

4.2.2. Елемент <frameelements>

След дефиницията на фрейма лексиконът представя неговите фреймови елементи, които за *work/die Arbeit/работа* са общо шест на брой – Employee, Employer, Position, Field, Task, Compensation. Те са групирани под етикет <frameelements>.

```
<frameelements>
  <frameelement name="...">
    <definition lang="..."> ... </definition>
  </frameelement>
</frameelements>
```

Синтаксисът към този елемент е:

```
<!ELEMENT frameelements (frameelement+)>
<!ELEMENT frameelement (definition+)>
<!ATTLIST frameelement
  name (Employee | Employer | Position | Field | Task | Compensation) #REQUIRED
>
```

Всеки фреймов елемент се описва и превежда на съответните езици във <frameelement>. За да се направи по-обхватно категоризиране, към този елемент са добавени атрибути – става дума за *name* със стойности Employee | Employer | Position | Field | Task | Compensation. Елементът <definition> е негов дъщерен и има атрибут *lang* със същите стойности, както бе описано в началото за друг елемент. Тъй като фреймовите елементи са важна част от цялата разработка, ще бъдат илюстрирани подробно чрез програмния код. Първо ще се представи дефиницията, а след това как изглежда в XML-кода.

➤ Employee

`<frameelement name="Employee"> ... </frameelement>`

Дефинициите на фреймовия елемент *Employee* са:

The person who works to complete some Task for Compensation.

`<definition lang="en">The person who works to complete some <fe href="Task">Task</fe> for <fe href="Compensation">Compensation</fe></definition>`

Die Person, die daran arbeitet, eine Aufgabe für Kompensation zu erledigen.

`<definition lang="de">Die Person, die daran arbeitet, eine <fe href="Task">Aufgabe</fe> für <fe href="Compensation">Kompensation</fe> zu erledigen</definition>`

Лицето, което работи, за да изпълни Задача за Възнаграждение.

`<definition lang="bg">Лицето, което работи, за да изпълни <fe href="Task">Задача</fe> за <fe href="Compensation">Възнаграждение </fe></definition>`

➤ Employer

`<frameelement name="Employer"> ... </frameelement>`

Дефинициите на фреймовия елемент *Employer* са:

The person or institution who gives Compensation to the Employee.

`<definition lang="en">The person or institution who gives <fe href="Compensation">Compensation</fe> to the <fe href="Employee">Employee</fe></definition>`

Die Person oder Institution, die dem Mitarbeiter eine Vergütung gibt.

`<definition lang="de">Die Person oder Institution, die dem <fe href="Employee">Mitarbeiter</fe> eine <fe href="Compensation">Vergütung</fe> gibt</definition>`

Лицето или институцията, която дава Възнаграждение на Служителя.

`<definition lang="bg">Лицето или институцията, която дава <fe href="Compensation">Възнаграждение</fe> на <fe href="Employee">Служителя</fe></definition>`

➤ Position

`<frameelement name="Position"> ... </frameelement>`

Дефинициите на фреймовия елемент *Position* са:

The label given to the kind of employment (e.g. as an engineer).

`<definition lang="en">The label given to the kind of employment (e.g. as an engineer)</definition>`

Das Label für die Art der Beschäftigung.

```
<definition lang="de">Das Label für die Art der Beschäftigung</definition>
```

Етикетът, който е поставен на вида заетост (напр.инженер).

```
<definition lang="bg">Етикетът, който е поставен на вида заетост (напр. инженер)</definition>
```

➤ Field

```
<frameelement name="Field"> ... </frameelement>
```

Дефинициите на фреймовия елемент *Field* са:

The Field in which the Employee works (e.g. finance, construction).

```
<definition lang="en">The <fe href="Field">Field</fe> in which the <fe href="Employee">Employee</fe> works (e.g. finance,construction)</definition>
```

Das Feld, in dem der Mitarbeiter arbeitet.

```
<definition lang="de">Das <fe href="Field">Feld</fe>, in dem der <fe href="Employee">Mitarbeiter</fe>arbeitet</definition>
```

Полето, в което Служителят работи (напр. финансиране, строителство).

```
<definition lang="bg"><fe href="Field">Полето</fe>, в което <fe href="Employee">Служителят</fe> работи (напр. финансиране, строителство)</definition>
```

➤ Task

```
<frameelement name="Task"> ... </frameelement>
```

Дефинициите на фреймовия елемент *Task* са:

The action the Employee must do for the Employer.

```
<definition lang="en">The action the <fe href="Employee">Employee</fe> must do for the <fe href="Employer">Employer</fe></definition>
```

Die Aktion, die der Mitarbeiter für den Arbeitgeber durchführen muss.

```
<definition lang="de">Die Aktion, die der <fe href="Employee">Mitarbeiter</fe> für den <fe href="Employer">Arbeitgeber</fe> durchführen muss</definition>
```

Действието, което Служителят трябва да извърши за Работодателя.

```
<definition lang="bg">Действието, което <fe href="Employee">Служителят</fe> трябва да извърши за <fe href="Employer">Работодателя</fe></definition>
```

➤ Compensation

```
<frameelement name="Compensation"> ... </frameelement>
```

Дефинициите на фреймовия елемент *Compensation* са:

The payment an Employee receives in exchange for work.

```
<definition lang="en">The payment an <fe href="Employee">Employee</fe>
receives in exchange for work</definition>
```

Die Zahlung, die ein Mitarbeiter im Austausch für die Arbeit erhält.

```
<definition lang="de">Die Zahlung, die ein <fe href="Employee">Mitarbeiter</fe>
im Austausch für die Arbeit erhält</definition>
```

Възнаграждението, което Служителят получава в замяна на работа.

```
<definition lang="bg"><fe href="Compensation">Възнаграждението</fe>, което
<fe href="Employee">Служителят</fe> получава в замяна на
работа</definition>
```

Вътре в елемента `<definition>` са интегрирани допълнителни елементи `<fe>`, които ще се срещат и на други места. Те са от съществено значение, защото с атрибутите си *href*, които също имат стойностите Employee | Employer | Position | Field | Task | Compensation, реферират директно към вече дефинираните фреймови елементи.

4.2.3 Елемент `<words>`

Посредством следващия елемент `<words>` се специфицират отделните лексикални единици, които са разпределени в дъщерните елементи `<word>`, представяйки индивидуално всяка лема. Този елемент сам по себе си съдържа атрибут *type*, който определя съответните единици към каква част на речта спадат. Само някои части на речта са представени, тъй като не всички извикват фрейма. Неговите стойности са: multi-word | adjective | verb | preposition | noun. Синтаксисът в дефиницията изглежда така:

```
<!ELEMENT words (word+)>
<!ELEMENT word (expr+, meaning+, grammar*, example_sentences,
templates,other_forms*)>
<!ATTLIST word
    type (multi-word | adjective | verb | preposition | noun) #REQUIRED
>
```

Лексикалните единици на фрейма са общо 38 на брой. Част от тях се повтарят на структурно равнище и поради тази причина няма бъдат разглеждани всички единици в разработката, а само тези, които са представители за всяка от частите на речта или имат различна структура, която да заслужава специално внимание. За да се изпълни задачата на паралелен ФреймНет, лингвистичната информация и

анотираните примери са представени на трите езика. Елементът <word> има сложна структура и съдържа множество от дъщерни елементи. Със следващата илюстрация се представя скелетът на елемента.

```
<word>
  <expr>...</expr>
  <meaning>...</meaning>
  <grammar>
    <...>
  </grammar>
  <example_sentences>
    <...>
  </example_sentences>
  <templates>
    <...>
  </templates>
  <other_forms>
    <...>
  </other_forms>
</word>
```

Разбира се, всеки елемент съдържа и свои допълнителни дъщерни елементи и атрибути, за да се представи едно изследване, което цели да е достатъчно изчерпателно и информативно. Всеки един от дъщерните елементи на <word> ще бъде разгледан в отделни подточки и ще се анализират особеностите им с помощта на примери от кода на изследването.

4.2.3.1. Елемент <expr>

Елементът <expr> представя графично лексикалните единици на фрейма. Той съдържа вече разгледания атрибут *lang* със задължителните стойности en | de | bg.

```
<!ELEMENT expr (#PCDATA)>
<!ATTLIST expr
  lang (en | de | bg) #REQUIRED
>
```

С негова помощ се визуализират единиците и на трите езика, като по този начин става видима разликата в броя на извикващите лемни. Някои лексикални единици съдържат повече лемни, други са с малък брой, тъй като на български език се намират подходящи синоними, които увеличават броя на думите, но по-конкретна

информация може да се открие в главата, свързана с лингвистичната информация за единиците. Всяка една от тях е отделена в самостоятелен елемент <expr>. Ето и няколко примера. При глагола *anstellen/to hire/назначавам* се забелязва точно тази тенденция:

<expr lang="de">anstellen</expr>
<expr lang="en">to hire</expr>
<expr lang="en">to employ</expr>
<expr lang="bg">назначавам</expr>
<expr lang="bg">наемам</expr>
<expr lang="bg">наема</expr>
<expr lang="bg">наемам на работа</expr>
<expr lang="bg">наема на работа</expr>
<expr lang="bg">вземам</expr>
<expr lang="bg">взема</expr>
<expr lang="bg">вземам на работа</expr>
<expr lang="bg">взема на работа</expr>

При съществителното име *die Arbeitslosigkeit/unemployment/безработица* тази тенденция не е налице, където се среща само по един вариант и на трите езика:

<expr lang="de">die Arbeitslosigkeit</expr>
<expr lang="en">unemployment</expr>
<expr lang="bg">безработица</expr>

Съществуват и единици като *einen Job an nehmen/to accept a job/започвам работа*, които се числят към категорията *multi-word*. Тези единици са разглеждани в българския лексикон под формата словосъчетания.

<expr lang="de">einen Job annehmen</expr>
<expr lang="de">eine Stelle annehmen</expr>
<expr lang="en">to accept a job</expr>
<expr lang="en">to take a job</expr>
<expr lang="bg">започвам работа</expr>
<expr lang="bg">постъпвам на работа</expr>
<expr lang="bg">приемам работа</expr>
<expr lang="bg">поемам нов пост</expr>

Ролята на извикващи думи във фрейма се изпълнява и от множество прилагателни имена, например:

<expr lang="de">arbeitslos</expr>
<expr lang="de">unemployed</expr>

<expr lang="de">безработен</expr>

<expr lang="bg">незает</expr>

<expr lang="bg">неангажиран</expr>

Един единствен предлог може да се обособи като лексикална единица във фрейма. Става дума за немския предлог *bei/at/при*:

<expr lang="de">bei (DAT)</expr>

<expr lang="en">at</expr>

<expr lang="bg">при</expr>

<expr lang="bg">в</expr>

<expr lang="bg">за</expr>

4.2.3.2. Елемент <meaning>

Този елемент представя тълковното значение на лексикалните единици. Както бе описано в началото дефинициите са заимствани от българския WordNet. И този елемент съдържа атрибут *lang* с вече познатите стойности за немски, английски и български език.

```
<!ELEMENT meaning (#PCDATA)>
```

```
<!ATTLIST meaning
```

```
    lang (en | de | bg)#REQUIRED
```

```
>
```

За илюстрация е избран произволен пример от кода. Става дума за лемата *das Gehalt/pay/заплата* със съответните дефиниции:

- pay

This noun corresponds closely to the English "salary" implying a regular payment to an Employee for work, often disbursed on a regular monthly scale. It is not used to refer to the amount paid for hourly work.

```
<meaning lang="en">This noun corresponds closely to the English "salary"
implying a payment to an Employee ... </meaning>
```

- заплата

Лемата е съпоставима с немския и английския вариант. По дефиниция означава парично възнаграждение за положен труд, като изразява възнаграждаването на работодателя за извършена работа. Работодател, служител, възнаграждение са ядрените елементи на съществителното име.

```
<meaning lang="bg">Лемата е съпоставима с немския и английския вариант.
```

```
По дефиниция означава парично възнаграждение </meaning>
```

За този елемент е предпочетено да се обърне внимание само на дефиниция на английски език, като се игнорира текст на немски. Същата тенденция е запазена и в БОЛ. Идеята на това начинание е дефиницията да може да се свърже и с тази от английския ФреймНет за повече яснота и по-лесно сравняване между електронните ресурси. Освен същинско описание на значението при елементите на български език се среща както информация относно някои граматически правила, така и данни за задължително и допустимо обкръжение, видимо от следващите примери с лексикалните единици *das Vorstellungsgespräch/interview/интервю* и *der Kollege/colleague/колега*.

- интервю

Събеседване или интервю, което е по-разпространената лема под влиянието английски език, показва взаимодействието между потенциалния работодател и потенциалния служител. Извършва се подбор за дадена позиция в съответното поле – служителят трябва да убеди и покаже качествата си пред работодателя. Работодателят от своя страна преценява дали служителят е в състояние да изпълнява служебните си задължения. Често по време на интервюто за работа се разисква и темата за възнаграждението. Аналогично на американския феномен и на български език се популяризира и интервю за работа по телефона (телефонно интервю/събеседване). По дефиниция интервюто означава разговор под формата на въпроси и отговори, обикновено насочен към събиране на информация по конкретна тема. Като основно обкръжение може да се приведат фреймовите елементи работодател и служител.

<meaning lang="bg">Събеседване или интервю, което е по-разпространената лема под влиянието английски език, показва взаимодействието между потенциалния работодател ... </meaning>

- колега

С идентично значение като в немски и английски език. Лемата има следната дефиниция: всеки от партньорите в работен екип по отношение на останалите членове. Слабото склонение (Schwache Deklination), характерно за немски, няма израз на български език. Що се отнася до граматиката се виждат известни разлики в членуването. На български език не е характерна употребата на определен и неопределен пред съществителното, както в немски (der, ein и т.н.). С единицата се изразява равнопоставеността между отделните членове на даден екип. Като единствен ядрен елемент се откроява служител.

`<meaning lang="bg">`С идентично значение като в немски и английски език. Лемата има следната дефиниция: всеки от партньорите в работен екип по отношение на останалите членове. Слабото склонение...`</meaning>`

4.2.3.3. Елемент `<grammar>`

Следващата част от елемента `<word>` описва всички граматически феномени, които са характерни за някои лексикални единици. Елементите `<grammar>` не са задължителен компонент, т.е. не се явяват във всяка единица, а само там, където е сметнато за необходимо да се въвеждат граматични правила. Запазени са изходните изследвания от проекта за немския онлайн лексикон (G-FOL), които се виждат в елемент `<grammar>` с атрибут `lang="en"`. Българските граматични правила са добавени отново в елемент `<grammar>`, но със стойност на атрибута `bg`. Този елемент има сложна структура с множество дъщерни елементи, което се вижда в дефиницията на XML-файла.

```
<!ELEMENT grammar (rule*, synonym*, neologism*, parts_of_speech*, prefix*,
verbs*, non_compositional*, idiomatic_expressions*, comparisons_adjectives*,
adjectives_verbs*, adjectives_past_participles*, adjectives_present_participles*,
nouns_word_building*, nouns_number*, nouns_article*, nouns_adjectives*,
abbreviation*, prepositions_dative?, verbs_separable_prefixes*,
reflexive_verbs_pronouns*, colloquial_speech*, table*)>
<!ATTLIST grammar
    lang (de | en | bg) #REQUIRED
>
```

Те се различават в зависимост от обекта на анализ. За илюстрация на данните от G-FOL и българския лексикон е взета граматическата информация от лексикалните единици *angestellt/employed/нает* и *anstellen/to hire/назначавам*. Например структурата на първата единица изглежда по следния начин:

```
<grammar>
  <rule> ... </rule>
  <adjectives_past_participles>
    <example> ... </example>
  </adjectives_past_participles>
  <adjectives_present_participles>
    <example> ... </example>
  </adjectives_present_participles>
  <table>
```



```
< ... >
</table>
</grammar>
```

Както вече стана ясно от DTD-файла елемент `<grammar>` съдържа множество дъщерни елементи и атрибут *lang* със стойностите `en` | `bg`, съответно за английските и българските граматични правила. В XML-кода данните се илюстрират по следния начин:

```
<grammar lang="en"> ... </grammar>
<grammar lang="bg"> ... </grammar>
```

- **Елемент `<rule>`**

Този елемент е важна част от изследването, тъй като представя граматичните правила от съответните езици. Не съдържа допълнителни елементи, а само текстова информация. Особеното при него е, че при необходимост може да се явява повече от един път.

```
<!ELEMENT rule (#PCDATA)>
```

За немската част от лексикона за *angestellt/employed/наем* правило гласи: In German (just as in English), the past participles of verbs (with the -ed ending in English) can be used as adjectives, known as "participial adjectives". Add an adjective ending when appropriate. Even a verb's present participle can be used as an adjective. This form of the verb is similar in meaning to English ing-forms, and is formed in German by adding a "d" (and an adjective ending, if necessary) to the infinitive form of the verb. Adjectives formed in this way apply to the type of frame element that would fill the subject role of the verb (e.g. *überraschend* applies to a Stimulus).

```
<rule>In German (just as in English), the past participles of verbs (with the
-ed ending in English) ... </rule>
```

За представителя на български език правилото изглежда по следния начин: Представките са словообразователни морфемии, които стоят пред корена, като придават на думата ново речниково значение.

```
<rule>Представките са словообразователни морфемии ... </rule>
```

- **Елементите `<adjectives_past_participles>` и `<adjectives_present_participles>`**

Тези елементи са характерни за данните от немския ФреймНет, като те представят по два примера за двете разновидности.

```
<adjectives_past_participles>
```

`<example>`Seine Mutter war enttäuscht, dass er bei der Prüfung durchgefallen ist. (His mother was disappointed that he failed the test.)`</example>`

`<example>`Die enttäuschte Mutter weint. (The disappointed mother cries.)`</example>`

`</adjectives_past_participles>`

`<adjectives_present_participles>`

`<example>`Das Ende der Geschichte war überraschend. (The end of the story was surprising.)`</example>`

`<example>`Das war ein überraschendes Ende. (That was a surprising ending.)`</example>`

`</adjectives_present_participles>`

Става ясно, че те също имат допълнителни дъщерни елементи `<example>` и `<rule>`. В избраните примери `<rule>` не е представен, но се среща на други места. В DTD те имат следната структура:

```
<!ELEMENT adjectives_past_participles (rule*, example+)>
```

```
<!ELEMENT adjectives_present_participles (rule*, example+)>
```

- **Елемент `<prefix>`**

Той принадлежи към примера от българската част на лексикона, като представя представките, с които се образуват глаголи.

`<prefix>`В българския език за представки се използват предимно предлозите – без, из, в (въ), въз, с (съ), по, под, при, пред, през, на, над, до, от, за, зад, раз и др.`</prefix>`

Не се числи към елементите със сложна структура, а представя само текстово съдържание.

```
<!ELEMENT prefix (#PCDATA)>
```

- **Елемент `<verbs>`**

Той също принадлежи към единицата *anstellen/to hire/назначавам*. По подобие на `<adjectives_past_participles>` и `<adjectives_present_participles>` съдържа два дъщерни елемента: `<rule>` и `<example>`, с които се въвежда правило и примери към него. В XML изглежда така:

`<verbs>`

`<rule>`Глаголи, образувани с представки: в-, до-, за-, из-, на-, над-, о-, от-, по-, под-, при-, пре-, про-, раз- и др.`</rule>`

<example>впиша, допиша, запиша, изпиша, напиша, надпиша, опиша,
отпиша, попиша, подпиша, припиша, препиша, пропиша,
разпиша</example>

</verbs>

В DTD се дефинира в тага:

```
<!ELEMENT verbs (rule*, example+)>
```

- **Елемент <table>**

Този елемент съдържа граматически правила или примери в табличен вид, с които да се представят данните.

```
<table>
```

```
<table_row>
```

```
<table_header>Subject of Verb</table_header>
```

```
<table_header>Direct Object</table_header>
```

```
<table_header>Adjective applies to:</table_header>
```

```
<table_header>Examples</table_header>
```

```
</table_row>
```

```
<table_row>
```

```
<table_data>Stimulus</table_data>
```

```
<table_data>Experiencer</table_data>
```

```
<table_data>Experiencer</table_data>
```

```
<table_data>aufgeregt (worked up), schockiert (shocked), enttäuscht  
(disappointed)</table_data>
```

```
</table_row>
```

```
<table_row>
```

```
<table_data>Experiencer</table_data>
```

```
<table_data>Content or Stimulus</table_data>
```

```
<table_data>Content or Stimulus</table_data>
```

```
<table_data>gefürchtet (feared), gehasst (hated), geliebt  
(loved)</table_data>
```

```
</table_row>
```

```
</table>
```

Сам по себе си също съдържа дъщерни елементи, които оформят вида на таблицата.

```
<!ELEMENT table (table_row+)>
```

```
<!ELEMENT table_row (table_header*, table_data*)>
```

```
<!ELEMENT table_header (#PCDATA)>
```

<!ELEMENT table_data (#PCDATA)>

Както става ясно елементите <grammar> имат различни дъщерни елементи в зависимост от това какви правила се описват, защото всеки от тях характеризира специфичен феномен от българската граматика. Освен гореспоменатите етикети в останалите <grammar> се срещат и други елементи. Подредбата им в следващите подточки не е йерархична и не следва структурата на кода, а по-скоро изразява принадлежност към съответната част на речта и е направен опит за логическа последователност, например всички глаголи са групирани в една подточка, съществителните в друга и т.н.

- <synonym> и <neologism>

В тези елементи се дефинират синонимите и неологизмите в български език.

- <parts_of_speech> и <colloquial_speech>

Тук се представят частите на речта и се обяснява какво представлява разговорната реч.

- <verbs>, <verbs_separable_prefixes> и <reflexive_verbs_pronouns>

Първият елемент характеризира категорията вид на глагола, вторият обяснява глаголите с делими представки, а третият – рефлексивните глаголи.

- <non_compositional> и <idiomatic_expressions>

Двата етикета привеждат данни относно словосъчетанията и тяхната устойчивост.

- <nouns_adjectives>, <adjectives_verbs> и <comparisons_adjectives>

Първите два елемента представят информация относно образуването на прилагателни имена, а <comparisons_adjectives> прави анализ на категорията степенуване.

- <nouns_word_building>, <nouns_number> и <nouns_article>

Тези елементи са свързани със съществителните имена, като разглеждат, от една страна, словообразуването, а от друга, образуване на множествено число и членуване.

- <abbreviation> и <prepositions_dative>

Двата елемента нямат общо помежду си, като единият анализира съкращенията, а другият въвежда в тематиката на предлозите.

По-конкретни данни относно граматичната информация ще бъдат представени в главата, свързана с лингвистичния анализ на разработката.

4.2.3.4. Елемент <example_sentences>

След единиците, дефинициите и граматичната информация се въвеждат аотираните примери в елемент `<example_sentences>`, който от своя страна съдържа дъщерни елементи `<example_gr>`.

```
<!ELEMENT example_sentences (example_gr+)>
```

В тази фаза всеки един пример е анализиран през погледа на немски, английски и български език. Всички примери имат еднаква структура, единствената разлика между лексикалните единици е в количеството на примерите, защото някои от тях илюстрират и с малък брой задължителното обкръжение, докато за други са необходими повече аотирани изречения. За онагледяване на данните ще се вземат примерите от лексикалната единица *die Arbeit/work/работа*. Тя ще изпълнява ролята на представител за останалите единици, защото всички те имат еднаква структура.

```
<example_sentences>
  <example_gr>
    <ex> ... <ex>
  </example_gr>
</example_sentences>
```

- **Елемент `<ex>`**

Всеки `<example_gr>` съдържа по три елемента `<ex>`, които представят примерите от лексикона. Аналогично на други елементи и при него е добавен атрибут *lang* със стойностите `de` | `en` | `bg`, за да може да се категоризират примерите към съответните езици. Във всеки елемент `<ex>` са добавени елементите `<fe>` със съответния атрибут *href*, чрез който анализирани лемите от примерите се свързват с фреймовите елементи на главния фрейм. Те също не са ограничени в количеството си, а зависят от броя на задължителните аргументи, които предикатът на изречението има.

```
<!ELEMENT example_gr (ex)+>
<!ATTLIST ex
  lang (de | en | bg) #REQUIRED
>
```

От единицата *die Arbeit/work/работа* ще бъдат представени част от аотираните примери, за да се илюстрира структурата на XML-кода, тъй като тя е еднаква за всички лексикални единици.

```
<example_gr>
  <ex lang="de"><fe href="Employee">Die junge Frau</fe> sprach von ihrer
  Arbeit <fe href="Position">als Journalistin.</fe></ex>
```

<ex lang="en"><fe href="Employee">The young woman</fe> spoke of her work <fe href="Position">as a journalist.</fe></ex>

<ex lang="bg"><fe href="Employee">Младата жена</fe> говореше за работата си <fe href="Position">като журналистка.</fe></ex>

</example_gr>

<example_gr>

<ex lang="de"><fe href="Employee">Er</fe> konnte keine Arbeit <fe href="Field">in seinem Beruf</fe> finden.</ex>

<ex lang="en"><fe href="Employee">He</fe> could find no work <fe href="Position">in his profession.</fe></ex>

<ex lang="bg"><fe href="Employee">Той</fe> не успя да си намери работа <fe href="Field">по професията си.</fe></ex>

</example_gr>

За следващите примери в разработката е представена алтернативна форма на визуализация чрез акцентирание върху елементите <ex>:

1) Aber die Arbeit in der Landwirtschaft ist wesentlich produktiver.

<ex lang="de">Aber die Arbeit <fe href="Field">in der Landwirtschaft</fe> ist wesentlich produktiver.</ex>

2) But work in agriculture is substantially more productive.

<ex lang="en">But work <fe href="Position">in agriculture</fe> is substantially more productive.</ex>

3) Но работата в селскостопанството е съществено по-продуктивна.

<ex lang="bg">Но работата <fe href="Field">в селскостопанството</fe> е съществено по-продуктивна.</ex>

4) Das war seine Art zu zeigen, dass er seine Arbeit für seine Kinder machte.

<ex lang="de">Das war seine Art zu zeigen, dass <fe href="Employee">er seine</fe> Arbeit für seine Kinder machte.</ex>

5) That was his way to show that he did his work for his children.

<ex lang="en">That was his way to show that <fe href="Employee">he</fe> did <fe href="Employee">his</fe> work for his children.</ex>

6) Това беше неговият начин да покаже, че той вършеше работата си само за децата.

<ex lang="bg">Това беше неговият начин да покаже, че <fe href="Employee">той</fe> вършеше на работата <fe href="Employee">си</fe> само за децата.</ex>

4.2.3.5. Елемент <templates>

Като последна задължителна част от елемента <word> може да се въведе компонента <templates> с дъщерен <template_gr>, който по подобие на <example_gr> също се състои от три елемента <temp>.

```
<templates>
  <template_gr>
    <temp> ... </temp>
  </template_gr>
</templates>
```

Елементът <temp> се характеризира с атрибут *lang* и съдържа вече споменати стойности за немски, английски и български. За DTD-кодировката е характерен следният синтаксис:

```
<!ELEMENT templates (template_gr+)>
<!ELEMENT template_gr (temp+)>
<!ELEMENT temp (#PCDATA | fe)*>
<!ATTLIST temp
  lang (de | en | bg) #REQUIRED
>
<!ELEMENT template (#PCDATA | fe)*>
```

В този компонент са анализирани задължителните и допустимите обкръжения посредством темплейти, които определят броя на аргументите и адюнктите. Както предходния елемент и този не е ограничен в количеството на темплейтите, а зависи от броя на задължителните елементи. В разработката е обърнато внимание само на аргументите, а адюнктите са оставени на заден план.

```
<template_gr>
  <temp lang="de"><fe href="Employer">EMPLOYER</fe> stellt <fe
href="Employee">EMPLOYEE</fe> ein.</temp>
  <temp lang="en"><fe href="Employer">EMPLOYER</fe> hires <fe
href="Employee">EMPLOYEE</fe> ein.</temp>
  <temp lang="bg"><fe href="Employer">EMPLOYER</fe> наема <fe
href="Employee">EMPLOYEE</fe></temp>
</template_gr>
```

За следващата група темплейти ще се навлезе в детайл по подобие на предишния компонент.

7) EMPLOYER stellt EMPLOYEE als POSITION ein.

```
<temp lang="de"><fe href="Employer">EMPLOYER</fe> stellt <fe href="Employee">EMPLOYEE</fe> als <fe href="Position">POSITION</fe> ein.</temp>
```

8) EMPLOYER hires EMPLOYEE as POSITION.

```
<temp lang="en"><fe href="Employer">EMPLOYER</fe> hires <fe href="Employee">EMPLOYEE</fe> as <fe href="Position">POSITION</fe> ein.</temp>
```

9) EMPLOYER наема EMPLOYEE като POSITION.

```
<temp lang="bg"><fe href="Employer">EMPLOYER</fe> наема <fe href="Employee">EMPLOYEE</fe> като <fe href="Position">POSITION.</fe></temp>
```

4.2.3.6. Елемент <other_forms>

На последна позиция във <word> се намира елементът <other_forms>, който не е задължителен компонент от кода, а се появява само на определени места. В немския ФреймНет съществува компонент *Alternate Forms*, от който са заимствани формите за немски език. Характерен е за лемите, които спадат към частите на речта съществително име и глагол.

```
<other_forms>
  <tense>
    <form> ... </form>
  </tense>
</other_forms>
```

При съществителните имена се въвеждат формите за множествено число на немски и български език, а при глаголите – формите за сегашно, минало свършено и минало несвършено време:

```
<tense lang="de">
  <form tense="Präsens">die Angebote</form>
</tense>
<tense lang="bg">
  <form tense="Präsens">предложения</form>
</tense>
<tense lang="de">
  <form tense="Präsens">Er jobbt</form>
  <form tense="Perfekt">Er hat gejobbt</form>
  <form tense="Präteritum">Er jobbte</form>
```



```

</tense>
<tense lang="bg">
  <form tense="Präsens">Той работи</form>
  <form tense="Perfekt">Той работил</form>
  <form tense="Präteritum">Той работи</form>
</tense>

```

Елементът <other_forms> има дъщерен елемент <tense> с атрибут *lang*, който има стойностите *de* и *bg*, посредством които се представят формите на немски и английски. Всеки елемент <tense> от своя страна разполага с дъщерните <form>, които също са разширени с атрибути. В този случай става дума за *tense* със следните стойности: Präsens | Perfekt | Präteritum.

```

<!ELEMENT other_forms (tense)+>
<!ELEMENT tense (form)+>
<!ATTLIST tense
  lang (de | en | bg) #REQUIRED
>
<!ELEMENT form (#PCDATA)>
<!ATTLIST form
  tense (Präsens | Präteritum | Perfekt) #REQUIRED
>

```

Като цяло не се обръща голямо внимание на тези форми от лингвистична гледна точка. Те са по-скоро с илюстративен и обогатителен характер.

4.3. Лингвистична информация

В тази глава детайлно ще се разгледат елементите <grammar>, тъй като те съставляват важна част от лингвистичната информация в онлайн речника. Ще се обърне внимание на всеки елемент по отделно, защото съдържат уникална информация за всяка единица от лексикона и са представени различни аспекти от граматиката. При подбора на граматичната информация са използвани специализирани ресурси за български език и граматика от интернет пространството, като за коректността им е правена справка и с граматиката на Л. Андрейчин (Андрейчин 1944). Направен е обстоен анализ за езиците немски и български, като правилата от английски език не са главен обект на изследването. В частност е направено сравнение между граматични правила на немски и английски в немския ФреймНет, като тази проблематика е запазена и в настоящата разработка. В английския ФреймНет този тип данни не са главен обект на изследване. Немският

проект за онлайн речник ги равнопоставя заедно с останалите компоненти и прави подробен анализ на граматиката. В семантико-синтактичния речник също се надгражда информацията от английския ФреймНет, като се въвежда граматичен клас, който описва допустимата парадигма от форми на глаголите. Българският онлайн лексикон разглежда глобално тематиката и въвежда всякакъв тип граматична информация не само за глаголите, но и за останалите части на речта.

4.3.1. Граматични правила

Този тип данни не се появяват във всяка лексикална единица, а само там, където е сметнато, че подобен тип информация е необходима. От една страна, става дума за единици, при които например граматичните правила на български се различават от тези на немски език, а от друга, за единици, при които в елемент <meaning> се разискват накратко отделни правила от българската граматика, обяснени след това по-обстойно. На тези феномени ще се обърне внимание в следващите редове при анализа на отделните единици с елементи <grammar>. По отношение на съдържанието на тези етикети те се различават по между си – всеки си има собствени елементи, които засягат даден аспект от българската граматика. В разработката може да се обособят две основни групи правила:

- правила, които показват различията между немската и българската граматика

В тази група са представени правила от немската граматика, които са взаимствани от немския онлайн лексикон, и правила от българската граматика, които служат като еквивалент на немските и описват аспекти, отнасящи се до конкретната единица. Представени са също така допълнителни особености, ако такива са налични за дадената лема. Чрез този тип данни може да се направи диференциация между двата езика на равнище граматика, да улесни разбирането и изучаването им, както и да послужи за по-задълбочени езиковедски изследвания.

- правила, които накратко са били споменати и описани в елемент <meaning>, след което са обяснени по-подробно в обособени елементи

Тук са илюстрирани само правила от българската граматика, без да се прави сравнение между двата езика. Идеята е да се надгради информацията от G-FOL и да се обогати онлайн платформата. Правилата се появяват при единици, за които е необходимо да се знаят по-конкретни данни, като целта е да се избегнат потенциални грешки и да се спомогне правилната употреба в езика.

На този етап ще се анализират единиците с граматична информация самостоятелно, за да се представи граматичното съдържание в българския онлайн речник. Трябва да се има предвид, че при някои лексикални единици правилата

съвпадат, защото има леми, които спадат към една и съща част на речта или се употребяват по идентичен начин. Тези единици ще бъдат разгледани под общ знаменател. Тази тенденция е запазена само при немските леми. При <grammar> елементите от българския лексикон е предпочетено да се въвеждат данните еднократно, а не да се повтарят за всеки елемент. При необходимост потребителите могат да извикват граматичните данни по всяко време.

4.3.2. Първа група

Към първата група спадат единиците:

- *angestellt/employed/наем*

Граматичната информация при тази единица се ограничава до употребата на причастия във функцията на прилагателни имена през погледа на немски и английски език. В елемент <rule> се въвеждат съответните правила, които са както следва. За немски език това е: In German (just as in English), the past participles of verbs (with the-ed ending in English) can be used as adjectives, known as "participial adjectives". Add an adjective ending when appropriate. Even a verb's present participle can be used as an adjective. This form of the verb is similar in meaning to English ing-forms, and is formed in German by adding a "d" (and an adjective ending, if necessary) to the infinitive form of the verb. Adjectives formed in this way apply to the type of frame element that would fill the subject role of the verb (e.g. überraschend applies to a Stimulus).

<rule>In German (just as in English), the past participles of verbs (with the -ed ending in English) can be ... </rule>

За българската част правилото е: Прилагателните имена означават признак (качество или свойство) на явленията, означени със съществителните имена: добър, щастлив, работлив, широк, сребърен, пясъчен, дълготраен. Прилагателните имена се делят на качествени и относителни. Качествените прилагателни изразяват признаци (качества или свойства) на явленията. Могат да се степенуват: хубав, най-хубавата жена, умен мъж, по-умно дете. Относителните прилагателни изразяват признаци (качества или свойства) на явленията по отношение към други явления. Те не се степенуват освен в преносна употреба: златен пръстен (*по-златен пръстен), дървена маса, най-дървената глава. Прилагателните имена имат членувани и нечленувани форми за мъжки, женски и среден род единствено и множествено число (<http://uraa.eu/BG-newspaper/>).

<rule>Прилагателните имена означават признак (качество или свойство) на явленията, означени със съществителните имена: добър ... </rule>

След този елемент се въвеждат <adjectives_past_participles> и <adjectives_present_participles>, които представят примери за илюстрация на правилата в дъщерни елементи <example>. Ето и няколко извадки:

<example>Seine Mutter war enttäuscht, dass er bei der Prüfung durchgefallen ist. (His mother was disappointed that he failed the test.)</example>

<example>Das Ende der Geschichte war überraschend. (The end of the story was surprising.)</example>

<example>писах - писан, гледах - гледан</example>

<example>Намалява дялът на нежелаещите да гласуват.</example>

В допълнение към българската част има конкретни правила, които обясняват причастията и са въведени в допълнителни етикети <rule>, част от <adjectives_past_participles> и <adjectives_present_participles>:

Миналото страдателно причастие се образува от аористната основа (основата на минало свършено време) на глаголите от свършен или несвършен вид + морфемата Н или Т. По-често се срещат с морфема Н, а причастия с морфемата Т се образуват от глаголи от I-во спрежение на НА. Друга особеност е, че при глаголи, чиято аористна основа завършва на ОХ, се вмята гласната Е между корена и морфемата Н или Т. При глаголи от II спрежение, които имат аористна основа ИХ, променят гласния И с Е. В този случай лемата се употребява като прилагателно – заедно със съществителното име. Тогава то е в служба на съгласувано определение.

<rule>Миналото страдателно причастие се образува от аористната основа (основата на минало свършено време) на ...</rule>

Сегашното деятелно причастие е нелична глаголна форма, която се изменя по род и число и може да се членува. То се образува само от глаголи от несвършен вид, а в изречението изпълнява службата на съгласувано определение.

<rule>Сегашното деятелно причастие е нелична глаголна форма, която се изменя по род и число и може ...</rule>

Към края на всеки <grammar> елемент са въведени таблици <table>, които да обобщат до тук представената информация в табличен вид. Те може да съдържат както граматични правила, така и примери.

- *arbeitslos/unemployed/безработен*

При тази единица се среща един от най-обемните елементи <grammar> за целия лексикон, като се акцентира върху употребата на прилагателните имена и тяхното степенуване. Изследванията от немския ФреймНет са запазени в лексикона под формата на няколко елемента.

Първоначално се явява <rule> със следната информация: There are two main ways to use adjectives in German that parallel the ways adjectives are used in English. These usages are illustrated in the table below. Predicate adjectives are part of a sentence's predicate, the part that states something about the subject. When used in this way, the adjective typically follows a verb like sein ("to be") or werden ("to become"), and appears in its most basic form. Attributive adjectives directly attribute a quality to a noun by appearing before it in the sentence; no verb comes between the adjective and the noun it describes. In attributive uses, endings are added to the adjectives. At a minimum, an attributive adjective in German gets an "e" at the end, although there are several possibilities. Adjective endings are difficult to master, so if you are in your first few years of study, the take-away here is that attributive adjectives get endings (an "e" or more), and predicate adjectives do not.

<rule>There are two main ways to use adjectives in German that parallel the ways adjectives are used in English. These ... </rule>

След това започват обяснения относно степенуването на прилагателните, което се анализира в елемент <comparisons_adjectives > и съдържа: In the Alternate Forms tab, you can see the comparative (e.g. gut - besser, "good" - "better") and superlative (e.g. gut – am besten, "good" - "the best") forms of an adjective. German and English are similar in their uses of comparative; both languages add an "-er" ending to make comparative forms, for example: wütend, wütender ("angry, angrier"), informiert, informierter ("informed, moreinformed"), etc. The main difference is that English sometimes does not allow such an ending (e.g. *stupider, *informeded, *loster), but in German, the "-er" ending is always possible, and "more" does appear with an adjective to convey the comparative meaning.

<comparisons_adjectives>In the Alternate Forms tab, you can see the comparative (e.g. gut - besser, "good" - "better") ... </comparisons_adjectives>

В <nouns_adjectives> се представят прилагателни, които са образувани от съществителни имена. Те са широко разпространени на немски език и при тях трябва да се спазват допълнителни правила, които са обяснени както следва: So-called "adjectival nouns" are derived from adjectives. These nouns behave like adjectives in that their endings change depending on the gender, case, and whether one uses a definite article (or der-word, e.g. der, dieser, welcher) or an indefinite article (or ein-word, e.g. eine, meine, ihre). To create an adjectival noun, simply use the adjective without the noun (with the appropriate article and ending as if the relevant noun were there). For example, if you want to refer to a good looking person, take an adjective like "schön" ("beautiful," "handsome"), as in "ein schöner Mann," and drop the noun; now you have "ein Schöner"

("a good looking man"). To refer to a woman, change the gender of the article accordingly: "eine schöne Frau" becomes "eine Schöne" ("a good-looking woman"). The endings you will most commonly need to refer to people using adjectival nouns are listed in the chart below.

<nouns_adjectives>So-called "adjectival nouns" are derived from adjectives.

These nouns behave like ... </nouns_adjectives>

Всички тези данни са подкрепени от две таблици, въведени чрез етикетите <table>. Що се отнася до информацията за българските лексикални единици, те не се различават на структурно ниво от своите немски еквиваленти, но на съдържателно се забелязват известни разлики при образуването и употребата на българските прилагателни. Правилото подкрепя тези твърдения: Прилагателното име обикновено не се употребява самостоятелно в речта, а в съчетание със съществителни имена. Заедно със съществителното то образува именно словосъчетание, което изразява едно понятие. Тази връзка води до голямата граматическа зависимост на прилагателното от съществителното. Прилагателното име няма свой собствен граматически род, няма и собствено граматическо число. То се мени по род и число, за да се съгласува с рода и числото на съществителното, към което се отнася. Най-честата и най-характерната употреба на прилагателните имена е атрибутивната – като съгласувано определение към някое съществително в изречението. Те се употребяват и като сказуемно определение (<http://znam.bg/com/>).

<rule>Прилагателното име обикновено не се употребява самостоятелно в речта, а в съчетание със съществителни имена ... </rule>

В елемента <nouns_adjectives> се добавят особеностите при образуването на съществителни имена, което също се различава от немското словообразуване: Като еквивалент на немските съществителни, които са образувани от прилагателни, на български език се явяват производните съществителни имена. Те са образувани от други думи, респ. от основи или корени на думи, посредством афикси, съединителни морфемии или по друг начин (стар + ец = старец, добър + ина = добрина); <http://uraa.eu/BG-newspaper/>.

<nouns_adjectives>Като еквивалент на немските съществителни, които са образувани от прилагателни ... </nouns_adjectives>

За разлика от немския ФреймНет при българския лексикон за тази единица не е сметнато за необходимо да се добавят таблици за илюстрация.

- *bei/at/npu*

За тази единица накратко са обяснени предлозите, които извикват дателен падеж. В <rule> се представя обща информация: For these prepositions, the noun that

follows must be in the dative case. The charts below show the articles in dative. Don't forget to add an extra "n" to dative plurals!

<rule>For these prepositions, the noun ... </rule>

В елемент <prepositions_dative> се представят всички предлози, които се използват със задължителен дателен падеж, а в таблица са илюстрирани формите на определителните и неопределителните членове в дателен падеж.

<prepositions_dative>aus, außer, bei, mit, nach, seit, von, zu</prepositions_dative>

Относно българската граматика е представена по обща информация, като са анализирани предлозите като цяло, а не в частност. В елемент <rule> се изписват правилата за предлозите, а след това в табличен вид се разграничават видовете предлози. Както се вижда от правилото има разлика при употребата им. Една от най-важните е липсата на задължителна обвързаност с конкретни падежи, което явление е характерно за немските предлози: Предлозите изразяват отношения между явленията, изразени с думи, които са самостоятелни части на речта: *без, в, във, вместо, въз, връз, върху, въпреки, до, за, зад, заради, из, извън, издад, измежду, изпод, като, край, към, между, на, над, накрай, наместо, насред, о, около, освен, от, отвъд, откъм, относно, отсам, оттам, оттатък, по, под, подир, покрай, помежду, поради, посред, посредством, пред, през, при, против, пряко, с, със, свръх, след, според, спрямо, сред, срещу, у, чрез*. По състав предлозите се делят на прости (съдържат само една морфема: *до, над, за*) и сложни (образуват се от свързването на простите предлози с други предлози, наречия или имена: *покрай, извън, свръх*). Някои предложни изрази са устойчиви словосъчетания: *с оглед на, по отношение на, въз основа на, подобно на, благодарение на, независимо от* и др. Предлозите *в* и *с* се удвояват пред думи, които започват съответно с В или Ф и С или З: *във влака, във фурната, със синя рокля, със злато* (<http://znam.bg/com/>; <http://gramatika-bg.com/>).

<rule>Предлозите изразяват отношения между явленията, изразени с думи, които са самостоятелни ... </rule>

- *der/die Angestellte/employee/служител;*
der/die Vorgesetzte/supervisor/ръководител;
die Arbeitslosigkeit/unemployment/безработица

Тези единици съдържат елемент <nouns_adjectives>, който вече е представен в единицата *arbeitslos*, където се анализират обстойно прилагателните имена с допълнителни правила и таблици. При необходимост може да се маркират и извикат тези граматически правила.

<nouns_adjectives>

<rule>So-called "adjectival nouns" are derived ... </rule>

</nouns_adjectives>

- *einstellen/to hire/назначавам*

В немския онлайн лексикон са представени глаголите с делими представки и техните особености при употребата, което на български език не е релевантно. В елемент <verbs_separable_prefixes> е видна цялата необходима информация. Чрез <rule> се привежда немското правило: Some verbs have a prefix that moves around in the sentence, depending on what form the verb takes. In the infinitive form, the prefix is attached, like "ausgehen" (to go out). If the verb is conjugated (in present or simple past tense), the prefix appears at the end of the clause, as in "Ich gehe heute Abend aus".

<verbs_separable_prefixes>

<rule>Some verbs have a prefix that moves ... </rule>

</verbs_separable_prefixes>

За българският еквивалент е въведен вече споменат елемент в разработката. Става дума за граматичната информация от единицата *bezahlen/to pay/плащам*, по-конкретно елементът <prefix>; източник на правилата е страницата <http://gramatika-bg.com/>.

<prefix>В българския език за представки се ... </prefix>

- *pensioniert/retired/пенсиониран*

Аналогично на предходните прилагателни имена и при тази единица се съдържат вече характеризирани елементи. Използвана е дефиницията от предходните единици в <rule>, <comparisons_adjectives>, <adjectives_past_participles>, <adjectives_present_participles>, като правилата са взаимствани от <http://znam.bg/com/>. Същите елементи са добавени в българската част на лексикона. Това, което заслужава по-специално внимание е един компонент, който се явява само при тази единица в немската част от лексикона, а именно елемента <adjectives_verbs>. При него са представени причастията на глаголите във функцията на прилагателни имена. Пояснява се словообразуването, изтъкнати са особеностите при употребата и т.н. Конкретни данни има в самия текст: In German (just as in English), the past participles of verbs (with the -ed ending in English) can be used as adjectives, known as "participial adjectives." Add an adjective ending when appropriate. Even a verb's present participle can be used as an adjective. This form of the verb is similar in meaning to English ing-forms, and is formed in German by adding a "d" (and an adjective ending, if necessary) to the infinitive form of the verb. Adjectives formed in this way apply to the type of frame element that would fill the subject role of the verb (e.g. überraschend applies to a Stimulus, and). Example: enttäuschen, überraschen (normal

use as verbs) Jens enttäuscht seine Mutter. (Jens disappoints his mother.) Das Ende der Geschichte überrascht Lena. (The end of the story surprises Lena.) Adjectives from Past Participles: Example: enttäuschen (to disappoint) > enttäuscht Seine Mutter war enttäuscht, dass er bei der Prüfung durchgefallen ist. (His mother was disappointed that he failed the test.) Die enttäuschte Mutter weint. (The disappointed mother cries.) The way frame elements are realized with the verb determine what the adjective can be used to describe. Details are given in the table below.

<adjectives_verbs>In German (just as in English), the past participles of verbs (with the -ed ending in English) can be ... </adjectives_verbs>

- *sich bewerben um/to apply for/кандидатствам за*

Като последен пример от тази група правила може да се обособи *кандидатствам за*. На немски език той се числи към групата на възвратните глаголи и в немската част са въведени правила за тях, както и възвратните местоимения в табличен вид. Самото правило е както следва: Verbs that are used reflexively often carry areciprocal meaning or the meaning that the subject is performing the action of the verb on themselves, although some abstract verbs are used reflexively without such meanings. In any case, the reflexive pronouns and word order are the same. As a general rule, the reflexive pronoun should appear just after the subject, although the V2 rule trumps this one, so in a basic sentence, you will find: subject, verb, reflexive (e.g. Er verliebt sich in Melanie, "He is falling in love with Melanie").

<reflexive_verbs_pronouns>Verbs that are used reflexively often carry a reciprocal meaning or the ... </reflexive_verbs_pronouns>

При българския глагол е добавена информация за възвратните глаголи, аналогично на немския еквивалент, въпреки че *кандидатствам* сам по себе си не е възвратен. Данните са въведени в същия елемент <reflexive_verbs_pronouns> и гласят: Глаголите, при които действието се възвръща върху извършителя, се делят на обикновени (невъзвратни) и възвратни. Обикновените глаголи означават действия, които са насочени не към вършителя, а към други лица и предмети (Ученикът мие чашата. Той счупи прозореца). Възвратните глаголи означават действия и състояния, извършвани от вършителя и връщащи се върху него. Те се образуват с помощта на кратките форми на възвратно лично местоимение се или си (Ученикът се мие. Той си счупи колелото); <http://uraa.eu/BG-newspaper/> и <http://znam.bg/com/>.

<reflexive_verbs_pronouns>Глаголите, при които действието се възвръща върху извършителя, се делят на ... </reflexive_verbs_pronouns>

4.3.3. Втора група

Към втората група се числят:

- *einen Job annehmen/to accept a job/започвам работа*

При тази единица са въведени данни, които характеризират словосъчетанията, този процес е представен в елемент <rule>: В български език се различават свободни и устойчиви словосъчетания. При свободните словосъчетания отделните думи, които ги образуват, запазват своето речниково значение. При устойчивите словосъчетания една или повече думи не пазят своето самостоятелно речниково значение, а самото словосъчетание придобива ново значение (<http://znam.bg/com/>).

<rule>В български език се различават свободни и устойчиви словосъчетания.

При свободните ... </rule>

В елемент <non_compositional> се описват устойчивите словосъчетания с отделно правило: Фразеологичните, или устойчивите словосъчетания, са предмет на лексикологията. Има и специализирани фразеологични речници, където се обяснява смисълът им, напр. придържам се към буквата на закона означава изисквам точно спазване на формалната страна на нещо – наредба, разпореждане и пр.; проявявам педантизъм, ходя по нервите на някого означава дразня, нервирам някого, ходя по гайдата на някого означава изпълнявам безропотно желанията на някого, постъпвам така, както той иска, слушам някого във всичко. Както се вижда тези изрази имат различен смисъл от преките значения на съставлящите ги думи. Фразеологичните словосъчетания могат да бъдат синонимни помежду си или да имат за синоним една или няколко думи. Като синоними може да се приведат подстрекавам, подбуждам и т.н.

<rule>Фразеологичните, или устойчивите словосъчетания, са предмет на лексикологията. Има и ... </rule>

В допълнителни елементи <example> се въвеждат илюстративни примери.

<example>пускам муха в главата на някого</example>

<example>подпалвам фитила на някого</example>

Анализират се също така и идиомите в <idiomatic_expressions>, които имат същата структура като предходния елемент. Правилото е: Идиомите са такива словосъчетания, при които няма никаква връзка между думите, които ги съставят.

<rule>Идиомите са такива ... </rule>

<example>Ни в клин, ни в ръкав.</example>

Към края на граматичния елемент се въвежда елемент <table>, който представя информацията в табличен вид.

- *einen Job kündigen/to quit a job/напускам работа*

При този пример се въвежда по-обща информация относно граматиката. Тъй като в дефиницията на единицата става дума за части на речта, като граматична информация са представени именно частите на речта, чиито данни може да послужат и при разглеждането на други лексикални единици. Първоначално има встъпителни думи в <rule>: Думите се делят на групи, наречени части на речта, обединени от най-общото в лексикалното си значение, от общи морфологични категории, синтактични функции и словообразователни типове (<http://znam.bg/com/>; <http://gramatika-bg.com/>).

<rule>Думите се делят на групи, наречени части ... </rule>

На следваща стъпка <parts_of_speech> се задълбочава информацията: Изменяемите части на речта могат да променят формата си и да изразяват различни морфологични категории (граматични значения). Такива са съществителното име, прилагателното име, числителното име, местоимението и глаголят. Неизменяемите части на речта не могат да променят формата си и да изразяват различни морфологични категории (граматични значения). Такива са наречието, предлогът, съюзът, частицата и междуметието. Самостояните части на речта могат да бъдат части на изречението. Такива са съществителното име, прилагателното име, числителното име, местоимението, глаголят и наречието. Несамостояните (служебните) части на речта изразяват различни морфологични или синтактични отношения между думите, съединяват ги или ги подчиняват. Такива са предлогът, съюзът и частицата. Специфични се наричат частите на речта, които не се определят като самостоятелни или несамостоятелни. Тук спада междуметието.

<parts_of_speech>Изменяемите части на речта могат да променят формата си и да изразяват ... </parts_of_speech>

В таблица се илюстрират подробно частите на речта, заедно с примери.

- *anstellen/to hire/назначавам*

Това е първият глагол, който се появява в лексикона и точно поради тази причина е предпочетено при него да се представи категорията вид, тъй като тя е от съществено значение за правилната употреба на глаголите в български език. Това се случва в елемент <verbs> с познатите <rule> и <example>. Правилото гласи: Лексикално-граматичната категория вид изразява начина, по който протича действието. Значенията на категорията са свършен вид и несвършен вид. Глаголите от свършен и несвършен вид са различни думи, а не форми на един и същ глагол. Свършеният вид представя действието в неговата цялост. Несвършеният вид представя действието в процеса на неговото протичане. Глаголите от свършен вид не

могат да се свързват с думи, които означават начало, продължение или край (<http://gramatika-bg.com/> и <http://znam.bg/com/>).

<rule>Лексикално-граматичната категория вид изразява начина, по който протича действието. Значенията ... </rule>

<example>бодна - бода</example>

<example>дам - давам</example>

- *bezahlen/to pay/плащам*

При тази единица са представени словообразователните морфемии за образуване на глаголи. Те се дефинират в <rule>: Представките са словообразователни морфемии, които стоят пред корена, като придават на думата ново речниково значение (<http://gramatika-bg.com/>).

<rule>Представките са словообразователни морфемии ... </rule>

Елемент <prefix> съдържа видовете представки: В българския език за представки се използват предимно предлозите – без, из, в (въ), въз, с (съ), по, под, при, пред, през, на, над, до, от, за, зад, раз и др.

<prefix>В българския език за представки ... </prefix>

Като последен компонент се добавя <verbs>, при който се описват представките за образуването на глаголи: Глаголи, образувани с представки: в-, до-, за-, из-, на-, над-, о-, от-, по-, под-, при-, пре-, про-, раз- и др.

<rule>Глаголи, образувани с представки ... </rule>

<example>впиша, допиша, запиша ...</example>

- *das Angebot/offer/оферта*

На този етап се описва накратко словообразуването на съществителни имена. Съществува само един елемент <nouns_word_building>: Има два вида съществителни имена. Прости съществителни се образуват по три начина: от глаголни корени (имам-имане), от прилагателни със словообразователни наставки (красив-красота), от съществителни със словообразователни наставки (род-родина). Сложни съществителни се образуват със съединителна морфема О или Е (земеделец, водомер) или без – директно свързване (Ботевград); <http://uraa.eu/BG-newspaper/>.

<nouns_word_building>Има два вида съществителни имена. Прости съществителни се образуват ... </nouns_word_building>

- *der Auftrag/order/поръчка*

Тъй като в дефиницията на тази лема става дума за множествено число, се въвеждат правила за образуването му на български език. <rule> съдържа: В основата на граматическата категория число лежи количествен признак на именуваните обекти. Съществителното име може да означава един обект или неопределено

множество от еднородни обекти. От формата за единствено число се образува формата на множественото число, като се прибави към нея окончание или се извърши замяна на окончание с друго такова (<http://znam.bg/com/>).

<rule>В основата на граматическата категория число лежи количествен признак на именуваните обекти ... </rule>

В <nouns_number> са представени окончанията за образуване на множествено число: Окончания за образуване на множествено число: -ове (столове), -и (внуци), -а (крака), -ища (пътища), -я (копия), -та (колиета), -ена (имена), -еса (чудеса)

<nouns_number>Окончания за образуване на множествено ... </nouns_number>

- *der Kollege/colleague/колега*

При единицата *колега* е разискана проблематиката на членуването, тъй като между членуването на немски и членуването на български език има известни разлики. Във вече познатия елемент <rule> се представя категорията определеност: Категорията определеност (членуване) е характерна за съществителното име. Изразява се с морфема (т.н. определителен член), която се поставя след съществителното име и се пише слято с него (за разлика от немски език); <http://bgmateriali.com/>.

<rule>Категорията определеност (членуване) ... </rule>

Фактическият анализ се състои в елемента <nouns_article>, който описва рода в българския език. С помощта на три допълнителни дъщерни елемента се разграничават трите рода. За мъжки род данните са най-много и са описани по следния начин: С пълен член -ът/-ят се членуват съществителните имена от мъжки род в единствено число, когато в изречението те изпълняват ролята на подлог или сказуемно определение (Градът се разраства. Това е градът /Героят победи. Това е победителят). С непълен член -а/-я се членуват съществителните имена от мъжки род, когато те са второстепенни части в изречението (Аз видях адвоката). Съществителни имена с завършек -тел, -ар, -й и лемите ден, кон, път, цар получават определителен член -я/-ят.

<masculine>С пълен член -ът/-ят се членуват ... </masculine>

С елементите <feminine> и <neutral> се допълват женски и среден род със съответните значения:

<feminine>Определителният член на съществителните в женски род, единствено число, е -та (къщата). При лемии, които завършват на -т, съгласният звук се удвоява (радостта).</feminine>

<neutral>Определителният член на съществителни от среден род в единствено число е -то (окото).</neutral>

- *der Lebenslauf/resume/автобиография*

За тази единица е характерен <grammar>, който се състои от данни относно съкращенията в български език. Този аспект е засегнат в лексикона, защото в дефиницията на единицата се представя съответното съкращение на лемата, въпреки че става дума за латинска дума. Това е елементът <abbreviation> със съдържание: Съществуват два вида съкращения - графични и лексикални. Графичните съкращения принадлежат само на писмения текст. При четене и говорене те не се произнасят, а се заместват с пълния вариант на съкратените думи. Лексикалните съкращения принадлежат и на писмената, и на устната реч. При четене и говорене те функционират като самостоятелни думи успоредно с пълния вариант. Графичните съкращения се използват при собствени имена, названия на обекти, титли и длъжности, мерни единици (П. Славейков, гр. Враца, д-р). Графичните съкращения задължително се пишат с малка буква и се означават с точка. Лексикалните съкращения представляват съкращения на имена, изрази или сложни думи. Най-разпространени са т. нар. абривиатури (ЕС, НДК); <http://ibl.bas.bg/>.

<abbreviation> Съществуват два вида съкращения - графични и лексикални.

Графичните съкращения принадлежат ... </abbreviation>

- *der Unternehmer/entrepreneur/предприемач*

За тази лема са характерни допълнителни синоними, които се срещат и при други лексикални единици. Поради тази причина накратко е представена дефиницията на синонимите в етикет <synonym>: Това са думи, които имат различен звуков състав, но са близки или еднакви по значение. Синонимите почти винаги се отличават един от друг или по своето лексикално значение, или по своята емоционално-експресивна окраска, или по стилистичната принадлежност към определен жанр на речта, или по своята употреба, или по способността да се свързват с други думи (<http://zamatura.eu/>).

<synonym> Това са думи, които имат различен звуков състав, но са близки или еднакви по значение ... </synonym>

- *die Arbeitsstelle/position/работно място*

При тази единица е въведена информация, която не реферира към конкретно граматично правило. Дефинират се неологизмите, за да се избегне объркване при категоризирането на единицата. В елемент <neologism> са въведени необходимите данни за изясняване на неологизмите: Неологизъм в езикознанието е дума, която се е появила наскоро в езика и която е приета в научни речници. Думи, които са употребени еднократно от някого или известно време в тесни среди (жаргон) не са неологизми и не се намират в речници. Неологизъм може да е освен една нова дума,

така и дума с ново значение, както и нова съставна дума, съкратена дума, акроним, криптограма и др. (<http://zamatura.eu/>).

<neologism>Неологизъм в езикознанието е дума, която се е появила наскоро в езика и която е приета ... </neologism>

- *jobben/to work/работя допълнително*

Като последна лема от втората група правила се обособява *jobben/to work/работя допълнително*, който съдържа сред преводните си варианти и разговорна дума *бачкам*, което е причина за кратко обяснение на разговорната реч. Това се случва в елемент <colloquial_speech>: Разговорна реч е езикът, който се употребява в устната реч при неофициален разговор, по време на диалог от ежедневието и въобще по време на вербална комуникация в тесен, роднински, приятелски или колегиален кръг (<http://bgmateriali.com/>).

<colloquial_speech>Разговорна реч е езикът, който се употребява в устната реч при неофициален разговор, по ... </colloquial_speech>

5. Обобщение на резултатите

В тази глава ще се направи кратко описание на получените резултати, което е категоризирано в три части.

5.1. Особености при елементите <expr>

Забелязва се, че лемите от българската част са в повечето случаи с по-голям брой от техните еквиваленти на немски и английски, тъй като българският онлайн лексикон се основава на синонимните множества от БулНет. За всяка дума е намерен съответният еквивалент на български, като са добавени и подходящи синоними, които също имат голяма честота на употреба за дадената ситуация.

За илюстрация може да се вземе единицата *angestellt/employed/заем*, при която на български език се срещат още лемите *нает/назначен/деен*. Като причина може да се изтъкне фактът, че една от задачите на изследването е да открие възможно най-много лексикални единици, които биват извиквани на български език. От друга страна, съществуват и единици, при които има само един възможен вариант на български. Такъв е случаят с *Arbeit suchen/to look for work/търся работа*, където не се срещат допълнителни лемите или съчетания с подобно значение. Ще бъдат

анализирани само тези лемни, които се различават по някакви критерии от техните еквиваленти, за да се направи сравнение между езиците.

Една от най-съществените разлики се забелязва още в началото при лексикалната единица *einen Job annehmen/to accept a job/започвам работа*. Докато на немски и английски присъства неопределителен член *einen/a*, на български език не е характерна неговата употреба. За голяма част от лексикалните единици на фрейма е характерна подобна закономерност. Става дума за категорията детерминация, където са включени три грамеми: определителен, неопределителен и нулев член. На немски език определителният член се използва, когато се говори за нещо познато и известно, а неопределителният – когато се говори за нещо непознато, неизвестно, неспоменавано (теорията за информационната структура). Артикли се пропускат само пред определени видове съществителни – събирателни и абстрактни съществителни, хранителни продукти и материали, роднински връзки и професии.

Често се допускат грешки при превод от немски на български, защото на немски език в голяма част от случаите трябва да се поставя определителен или неопределителен член пред съществителните имена, докато на български не винаги е задължителен. За да се изясни защо се стига до такива пропуски на български език, в тази си част изследването се основава на граматиката на Р. Ницолова (Ницолова 2008). Тя установява, че най-често срещан в текст е определителният член, а нулевият член и *един* се употребяват по-рядко, като изразът с *един* може да се срещне само веднъж, когато за пръв път се представя неговият референт. Само определителният член служи като средство за означаване на кореферентност. В семантичен план определеността е вид характеристика както на екстензията на името, т.е. количеството обекти, така и на интензията му, т.е. съдържанието (Ницолова). Употребата на артиклите може да се класифицира на два вида: референтна и нереферентна. При първия вид в реалния свят съществува предмет, който се съотнася с израз с член. В семантичната структура изразът е аргумент или адюнкт към предиката. При нереферентната употреба името с нулев член или с *един* означава само пропозиционална функция *da si x* и е предикат. С единицата *започвам работа* ще се представи детерминацията на български език. Употребата на нулев член подсказва за съществуваща работа от сферата на труда, като говорещият се абстрахира от специфичните признаци на дейността. При *започвам една работа* става дума за съществуваща работа, като се причисляват някакви индивидуални признаци. Когато се използва определителен член *започвам работата* се посочва, че работа се числи към сферата на труда, обозначават се някакви индивидуални

признаци и характеристики, които са познати както на говорещия, така и на слушателя. В този случай слушателят може да определи еднозначно референта.

При единицата *anstellen/to hire/назначавам* се наблюдава нещо друго: сред вариантите на български език освен глаголите *назначавам/наемам/наема/вземам/взема* могат да се обособят и устойчиви съчетания от съществително и глагол, като например *наемам на работа/наема на работа/вземам на работа/взема на работа*. В този случай съществителното *работа* е извикано от глаголите, т.е. намира се в тяхната синтактична рамка. Друга особеност е, че немският глагол се числи към типа глаголи с делими представки за разлика от английски и български. Когато глаголът е в просто изречение, представката се отделя от основата и се поставя в края на изречението. Ударението пада върху представката.

При *Arbeit suchen/to look for work/търся работа* употребата на различни предлози извиква отделните фреймови елементи. Единствената разлика се наблюдава при английския еквивалент, където се среща предлога *for*, който е неразделна част от глагола *look* и образува съчетание с ново значение: *to look for*. В немски език също съществува рекция, т.е. задължителна обвързаност между глагол/съществително/прилагателно/наречие → определен предлог → определен падеж → определено значение.

Случаят с единицата *arbeiten/to work/работя* също заслужава кратък анализ, тъй като съдържа допълнителен глагол. Става дума за *занимавам се* – възвратен глагол с друго значение, но в определени ситуации може да се използва и със значение на *работя, трудя се*. Немските възвратни глаголи са съпроводени винаги от възвратно местоимение *sich*.

На български език категорията вид на глагола има съществено значение. При голяма част от единиците с глаголи са въведени двата вида глаголно действие. За илюстрация се привежда *bezahlen/to pay/плащам. Плащам/заплащам* е от несвършен, докато *платя/заплатя* – от свършен вид. Тази категория не е характерна за езиците немски и английски. В граматичната част на изследването е обърнато специално внимание на вида, с допълнителни правила и примери.

За съществителните имена от немската част е характерен задължителен определен член. В зависимост от рода се различават три вида – *der/die/das (das Angebot, das Gehalt)*. Той е неразделна част от немските съществителни. По-конкретни данни за определеността се намират в началото на тази глава (Ницолова 2008) и в частта с граматичните правила.

При *die Stellenanzeige* става дума за сложно съществително име, което няма конкретен превод на български език, а се представя описателно чрез *обява за работа*. Подобна е ситуацията и с английския вариант – *job ad*.

Немското словосъчетание *in Rente gehen* на български език може да се представи също чрез словосъчетание *излизам в пенсия*, но по-разпространен е глаголят *пенсионирам се*.

Сред последните единици в изследването се откроява *sich bewerben um/to apply for/кандидатствам за*, където се описва употребата на възвратни глаголи. Несъответствия се откриват в българските лема, защото има преводен вариант, който не е възвратен глагол.

Последната единица от фрейма е *jobben*. На български език няма конкретен превод, а се представя описателно чрез допълнително прилагателно, за да се изрази правилното значение на глагола: *работя допълнително*.

5.2. Особености при елементите <grammar>

В тази част ще се разгледат особеностите при граматичните правила на немски и английски език. Акцентът е поставен върху прилагателните имена, защото те са сред най-многобройните единици, при които са въведени правила от граматиката на двата езика. Ще се открият прилики и разлики при употребата им, както и допълнителни специфики, характерни за формообразуването.

Първата лема, при която се сравняват правила от немски и български, е прилагателното *angestellt/employed/заем*, където са представени причастията и тяхната употреба. Както на немски, така и на български език миналите страдателни причастия могат да се употребяват и като прилагателни. При българските форми са характерни две морфеми за формообразуването – Н и Т. При немските варианти се използва формата Partizip II (миналото причастие на глагола). При употребата като съгласувано определение прилагателното се склонява и се добавят съответните окончания в зависимост от рода на съществителното, а когато е използвано като сказуемно определение – формата остава непроменена. При сегашното деятелно причастие ситуацията е идентична, добавят се съответните морфеми, които са D (немски) и Щ (български). Употребява се също като съгласувано и сказуемно определение.

При единицата *arbeitslos/unemployed/безработен* също е обърнато внимание на предикативната и атрибутивната употреба на прилагателните. Но освен това се прави анализ и на степенуването, което се различава до известна степен между езиците. На немски и английски се степенува с помощта на допълнителни суфикси

ER и (E)ST, докато на български формите изглеждат по друг начин. Степенува се с помощта на частиците ПО за сравнителна и НАЙ за превъзходна степен, като между тях и формата на прилагателното се поставя малко тире. В самото изследване степенуването е илюстрирано в табличен вид. В тази единица като граматическа информация са представени и съществителните, образувани от прилагателни. Такъв тип думи не са широко разпространени на български език – като техен еквивалент са приведени производните съществителни, които са образувани посредством афикси (добър + ина = добрина).

Единственият представител на частта на речта предлог е единицата *bei/at/npu*. Като разлика между предлозите в немски и български може да се обособи задължителната обвързаност с конкретен падеж, тъй като немският предлог *bei* се употребява с дателен падеж, докато падежните форми в български език се считат за остарели и не се употребяват.

5.3. Особености при елементите <example_gr> и <template_gr>

Тези части от изследването ще бъдат разгледани под общ знаменател, тъй като се проследяват еднакви закономерности и при двата компонента.

Още при първата единица *einen Job annehmen/to accept a job/започвам работа* се среща вече разискваната тематика относно поставянето на определителен и неопределителен член. Със съответните примери и темплейти може да се илюстрира специфичността, респективно неспецифичността на именните изрази при употребата на артиклите. Тази проблематика се среща и при други единици в изследването:

Die Fotografin nimmt einen Job in einem Verlag an./EMPLOYEE nimmt einen Job an.

The photographer is taking a job in a publishing house./ EMPLOYEE accepts a job.

Фотографът започна работа в издателство./ EMPLOYEE започна работа.

При следващата единица *einen Job kündigen/to terminate employment/напускам работа* се забелязва друга закономерност. Става дума за присъствието на лични местоимения в изречението. Докато на немски употребата на местоименията е задължителна, на български език тяхното присъствие се смята за редундантно, т.е. излишно. Въпреки това те са запазени при аотираните изречения в българската част, за да се илюстрират възможно най-подробно фреймовите елементи към съответните единици.

Ich kündige meine Stelle./EMPLOYEE kündigt.

I am quitting my job./ EMPLOYEE quits.

Аз напускам работното си място./ EMPLOYEE напуска.

Чрез употребата на глагола *напускам работа/служба* на български език се неутрализира характерният израз за немски и английски *прекратявам договор с нкг*.

При *angestellt/employed/заем* се проследяват известни разлики в словоредата, тъй като в немски той се подчинява на глаголната форма, което в случая представлява рамкова конструкция със спомагателен глагол на второ място, а пълнозначният глагол се поставя в края на изречението:

EMPLOYEE ist als POSITION angestellt.

EMPLOYEE is employed as POSITION.

EMPLOYEE е нает като POSITION.

При единицата *anstellen/to hire/назначавам* се откроява друга разлика между словоредата на немски и български, тук става дума за вече обсъждана проблематика, а именно глаголите с делими представки. Този тип е характерен за немски език, където представката се разделя от глагола и се поставя на последна позиция в изречението, изглежда както следва:

Die Firma stellt jeden Oktober 20 Kassierer an./EMPLOYER stellt EMPLOYEE an.

The firm hires every October 20 cashiers./EMPLOYER hires EMPLOYEE.

Фирмата наема на работа 20 касиери всеки октомври./ EMPLOYER наема EMPLOYEE.

При единицата *bezahlen/to pay/плащам* се забелязва друго несъответствие. От една страна, немският глагол се употребява със задължителен винителен падеж, а от друга, той се числи към групата глаголи с неделими представки. На български език плащам се употребява с предлог на (плащам – На кого). Избраният пример е в страдателен залог:

EMPLOYEE wird bezahlt.

EMPLOYEE is payed.

На EMPLOYEE се плаща.

При *заплащам* например не е необходима употребата на допълнителен предлог на, както при предходния глагол. Когато се използва с предлог *за*, е в значение на *давам пари за извършена работа*.

Die Jugendlichen werden für ihre Arbeit bezahlt.

The youths are paid for their work.

Младежите биват заплащани за работата си.

За разлика от немски и английски, при които принадлежност и собственост се изразява с морфемата *S* (при английски се поставя апостроф като разделителен

символ), на български това се случва посредством предлог, видимо от единицата *das Gehalt/pay/заплата*:

EMPLOYEEs Gehalt.

EMPLOYEE's salary.

Заплатата на EMPLOYEE.

Инфинитивните конструкции в немски *zu* + пълнозначен глагол се представят на български език с *да*-изречения, което служи за въвеждане на фреймов елемент TASK. Тази особеност се среща при съществителното *der Auftrag/order/поръчка*.

1939 bekam er den Auftrag, das neue Parlamentsgebäude in Ankara zu bauen./EMPLOYER gibt EMPLOYEE den Auftrag, TASK (Infinitivsatz).

In 1939 he got the contract to build the new parliament building in Ankara./EMPLOYER gives EMPLOYEE the order, TASK (infinitive clause).

През 1939 г. той получава поръчката да построи новата сграда на парламента в Анкара./EMPLOYER възлага на EMPLOYEE поръчка, TASK (да-изречение).

Както стана ясно от *bei/at/npu*, немските предлози в изследването изискват употребата на конкретен падеж. Тази зависимост се среща и при други примери, където на определени места се изисква употребата на падежи в немски. Съответния падеж на съществителното е отбелязан в скоби:

EMPLOYEE arbeitet im Auftrag EMPLOYER (genitive).

EMPLOYEE works under contract of EMPLOYER.

EMPLOYEE работи по поръчка на EMPLOYER.

В края на изследването единицата *pensioniert/retired/пенсионирани* също заслужава внимание, защото при нея се откриват разлики в атрибутивната употреба. За немски и български език е характерна употребата на допълнителни суфикси, защото миналото страдателно причастие се склонява, докато на английски формата не се съгласува със съществителното и остава непроменена.

Der pensionierte Allgemeinarzt kommt aus Köthen in Sachsen-Anhalt.

The retired general doctor comes from Köthen in Sachsen-Anhalt.

Пенсионираният личен лекар идва от Кьотен, Саксония-Анхалт.

С последното несъответствие се изчерпват основните разлики, които може да се открият в българските и немските единици, които са свързани с фрейма *work*.

6. Заключение

Българският онлайн лексикон може да се обозначи като триезичен паралелен ФреймНет. Създаден е по подобие на многоезичната лексикална база от данни (Боас 2005), която използва семантичните фреймове от английския ФреймНет. От една страна, БОЛ съдържа голямо количество от допустими семантични и синтактични комбинации, а от друга, за всяка лексикална единица се предлагат многобройни аотирани примери. Корпусът от примерите цели да отрази съвременното състояние на анализирани езици, като се представя семантичната структура и лексикалните значения на отделните единици. Изразяват се също така особеностите при синтактичната реализация на аргументите и честотата на употреба. Като съществени разлики между езиците може да се обособят:

- разминавания при употреба на артикли (детерминация);
- разлики при употребата на местоименията;
- липсата на обвързаност с конкретни предлози и падежи;
- наличието на устойчиви словосъчетания;
- разлики при възвратните глаголи;
- значението на граматическата категория вид;
- предикативна и атрибутивна употреба на причастията;
- склонение на прилагателни имена и въвеждане на допълнителни суфикси;
- несъответствия при степенуването на прилагателни;
- отклонения в словоредата на изречението;
- формообразуване – делими и неделими представки;
- инфинитивни конструкции и *да*-изречения.

Лексиконът може да се използва както за теоретични изследвания, така и при изучаването на някои от представените езици. Тъй като кодът на проекта е отворен, би могъл да послужи и при създаването на различни програми за компютърна обработка на естествения език.

7. Приложения

Тъй като приложенията са твърде обемни и съдържат основно програмен код и XML файлове, тук ще бъдат дадени само връзки към съответните интернет страници, където могат да бъдат разгледани и свалени.

7.1. Приложение 1

Български онлайн лексикон – връзки към проекта.

Уебстраница на проекта:

<http://www.ezik.bg/exist/apps/FrameNet>

Основният XML-файл на системата:

<http://www.ezik.bg/exist/apps/FrameNet/data/framenets/work.xml>

Основният DTD-файл на системата:

<http://www.ezik.bg/exist/apps/FrameNet/data/framenet.dtd>

7.2. Приложение 2

Диск с основните данни на проекта.

8. Библиография

- Андрейчин 1944: Андрейчин, Л. *Основна българска граматика*. София, 1944.
- Атанасов 2006а: Атанасов, А. Създаване на учебни граматически електронни речници. В: *Втора национална конференция с международно участие по електронно Обучение във висшето образование*. Китен, 14-17.09.2006, 32-35.
- Атанасов 2016: Атанасов, А. *Предложни фрази в предикативна позиция (аргументна структура)*. София 2016.
- Атанасов 2016: Atanasov, A. SynTags – web interface for syntactic and semantic annotation. In: *Proceedings of Second International Conference on Computational Linguistics in Bulgaria*, Sofia, 2016, 47-53.
- Бейкър и др. 1998: Baker, Collin F., Charles J. Fillmore, and John B. Lowe. *The Berkeley Framenet Project*. in COLING-ACL '98 Proceedings of the Conference held August 10-14, 1998, in Montreal, Canada, pp 86-90.
- Боас 2005: Boas, Hans C. Semantic Frames as Intelingual Representations for Multilingual Lexical Databases. *International Journal of Lexicography* 18(4), pp. 445-478.
- Бъркалова 1997: Бъркалова, П. *Българският синтаксис – познат и непознат. Увод в курса по синтаксис на СБЕ*. Пловдив, Университетско издателство.
- Бояджиев и др. 1998: Бояджиев, Т. И. Куцаров, Й. Пенчев. *Граматика на СБКЕ*. София, ИК „Петър Берон”.
- Коева 1999: Коева, С. *Синтаксис и пунктуация*. Пловдив.
- Коева 2001а: Коева, С. Малки изречения – *Съвременни синтактични теории*. *Помагало по синтаксис*, съставител С. Коева, Пловдив, Пловдивско университетско издателство, 61-68.
- Коева 2001б: Коева, С. Типология на подчинените изречения – *Съвременни*

- синтактични теории. Помагало по синтаксис*, съставител С. Коева, Пловдив, Пловдивско университетско издателство, 61-68.
- Коева 2001в: Коева, *Съвременни синтактични теории. Помагало по синтаксис*, съставител С. Коева, Пловдив, Пловдивско университетско издателство.
- Коева 2005б: Коева, С. *Аргументна структура. Проблеми на простото и сложното изречение*, СемаРШ, София 2005.
- Коева и др. 2008: Коева, С. Р. Влахова, Р. Декова, П. Несторова, А. Атанасов. *Българският ФреймНет. Семантико-синтактичен речник на българския език*. София 2008.
- Ландау 1989: Landau, S. I. *Dictionaries, the art and craft of lexicography*. Cambridge: Cambridge University Press.
- Ницолова 2008: Ницолова, Р. *Българска граматика. Морфология*. София, УИ „СВ. Климент Охридски“.
- Пенчев 1984: Пенчев, Й. *Строеж на българското изречение*. София.
- Пенчев 1993: Пенчев, Й. *Български синтаксис. Управление и свързване*. Пловдивско университетско издателство „Паисий Хилендарски“.
- Рупенхофър 2016: Ruppenhofer, J., Ellsworth, M., Petruck, M.R.L., Johnson C.R. and Scheffczyk, J. *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, California.
- Филмор 1982: Fillmore, C. J., Frame Semantics. In *Linguistics in the Morning Calm*, ed. by The Linguistic Society of Korea, Seoul: Hanshin, pp. 111- 137.
- Филмор 2002: Fillmore, C. J. FrameNet and the Linking between Semantic and Syntactic Relations. In Tseng, S.-C., editor, *Proceedings of COLING 2002*.

Интернет ресурси:

- БАН: Речник на българския език: <http://ibl.bas.bg/rbe/>
- БНК: Български национален корпус: <http://search.dcl.bas.bg/>
- BulNet: Българският WordNet: <http://dcl.bas.bg/bulnet/>
- Duden: Немски тълковен речник: <http://www.duden.de/>
- Framenet: Английският ФреймНет: <http://framenet.icsi.berkeley.edu/fndrupal/>
- Framenet: Немският ФреймНет: <http://www.laits.utexas.edu/gframenet/>
- G-FOL: Немският онлайн лексикон: <http://coerll.utexas.edu/frames/>
- 3WSchools: Интернет наръчници: <http://www.w3schools.com/>

Страници с граматични правила:

<http://uraa.eu/BG-newspaper/>

<http://bgmateriali.com/>

<http://zamatura.eu/>

<http://znam.bg/com/>

<http://gramatika-bg.com/>

<http://ibl.bas.bg/>